Copy No. _____

Defence Research and Development Canada
Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

Development of Measures of Effectiveness and Performance from Cognitive Work Analysis Products

Gerald Lai and Tab Lamoureux
CAE Professional Services

Prepared by:
CAE Professional Services
1135 Innovation Drive
Ottawa, ON K2K 3G7

Contractor's Document Number: 5254-001
Contract Project Manager: Matthew Keown, 613-247-0342
PWGSC Contract Number: W7707-098218/001/HAL
CSA: Mark Hazen, Group Leader, Maritime C2 Concepts Development, 902-426-3100

Defence R&D Canada – Atlantic

Contract Report
DRDC Atlantic CR 2011-282
February 2012

Canada

This page intentionally left blank.

Development of Measures of Effectiveness and Performance from Cognitive Work Analysis Products

Gerald Lai
Tab Lamoureux
CAE Professional Services

Prepared By:
CAE Professional Services
1135 Innovation Drive
Ottawa, ON Canada K2K 3G7

Contractor's Document Number: 5254-001
Contract Project Manager: Matthew Keown, 613-247-0342
PWGSC Contract Number: W7707-098218/001/HAL
CSA: Mark Hazen, Group Leader, Maritime C2 Concepts Development, 902-426-3100

Defence R&D Canada – Atlantic

Contract Report
DRDC Atlantic CR 2011-282
February 2012

Principal Author

Original signed by Gerald Lai

Gerald Lai

Consultant

Approved by

Original signed by Francine Desharnais

Francine Desharnais

Head, Maritime Information and Combat System Section

Approved for release by

Original signed by Leon Cheng

Leon Cheng

Chair, Document Review Panel

In conducting the research described in this report, the investigators adhered to the policies and procedures set out in the Tri-Council Policy Statement: Ethical conduct for research involving humans, National Council on Ethics in Human Research, Ottawa, 1998 as issued jointly by the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada.

Abstract

Defence Research and Development Canada (DRDC) – Atlantic is currently investigating new systems to support the VICTORIA Class Submarine command team. The program will include human-in-the-loop experimentation within a virtual environment facility known as vVictoria. The project had previously utilized human factors engineering processes such as Cognitive Work Analysis (CWA) to characterize the command team's work environment in support of new system design. This report investigated the utility of re-purposing the results of the design-based work to assist in developing measures of effectiveness (MOE) and performance (MOP) to support project experimentation. It was found that while the design work provided a lot of information, specific augmentation for measure development was required. Further, it was found that the addition of specific mission vignettes was required to reduce the numbers of measures coming out of the design work to a manageable level. This report details the development process, results of extra knowledge elicitation and validation, and resulting MOE/MOPs.

Résumé

Recherche et développement pour la défense Canada (RDDC) – Atlantique examine actuellement de nouveaux systèmes servant à soutenir l'équipe de commandement des sous-marins de la classe VICTORIA. Ce programme comprendra des essais avec intervention humaine dans le simulateur virtuel vVictoria. Dans le cadre de ce projet, des processus d'ingénierie des facteurs humains tels que l'analyse du travail cognitif (ATC) ont servi à caractériser le milieu de travail de l'équipe de commandement afin d'appuyer le nouveau concept du système. Ce rapport vise à déterminer si les résultats du travail de conception peuvent aider à élaborer des mesures de l'efficacité (MOE) et du rendement (MOP) en vue de réaliser les essais du projet. Beaucoup d'information a été recueillie durant le travail de conception, mais il faut davantage de renseignements spécifiques pour établir les mesures. En outre, d'autres scénarios de mission spécifique sont nécessaires pour amener le nombre de mesures obtenues avec le travail de conception à un niveau raisonnable. Le présent rapport porte sur le processus d'élaboration, les résultats liés à l'acquisition et la validation d'autres connaissances, ainsi que les MOE et les MOP qui en découlent.

This page intentionally left blank.

Executive summary

Development of Measures of Effectiveness and Performance from Cognitive Work Analysis Products

Gerald Lai; Tab Lamoureux; DRDC Atlantic CR 2011-282; Defence R&D Canada – Atlantic; February 2012.

Introduction: Defence Research and Development Canada (DRDC) – Atlantic have been tasked with the investigation of new systems for Command and Control (C2) aboard the VICTORIA Class Submarine. As part of this effort, DRDC Atlantic has used human factors (HF) techniques such as Cognitive Work Analysis (CWA) and Goal Directed Task Analysis (GDTA) to investigate the work carried out by the VICTORIA Class command team. The work has culminated in the design of an Information Integration Display (IID) which consolidates a number of critical information sources at one display. DRDC Atlantic has also constructed an experimentation facility, called the "Virtual VICTORIA", or "vVictoria", which will be used to conduct experiments to establish the usability and utility of tools such as the IID.

The current contract had two objectives: the development of Measures of Performance (MOPs) and Measures of Effectiveness (MOEs) for the evaluation of C2 activities and to investigate if work analyses performed for design purposes can be repurposed for the development of MOPs/MOEs.

Results: Through a review of the design based work analysis results, a total of 168 MOPs and MOEs were identified. These MOPS and MOEs were grouped according to nine functional areas and presented to operational submariners for validation, review and refinement. The functional areas were: Communications (external); Shared Situational Awareness (SA) (includes internal crew interaction); Safety; Ship Handling; Covertness; Planning; Contact Management; Individual SA; and Submarine Systems. The MOPs, MOEs, and the functional areas were considered to be representative and valid measurements of the activities performed aboard the VICTORIA Class Submarine. Following the validation exercise the MOP/MOEs were further categorized by developing a set of four scenario vignettes corresponding to particular primary mission types: navigation, contact management, planning ("window of opportunity"), and incident response. A subset of measures was chosen for each vignette which maximized coverage of the functional areas. Additionally, MOPs and MOEs were developed for each vignette for the IID.

Significance: The selection of robust informative measures is a critical aspect of all development and experimentation. This work has developed a significant number of potential MOPs and MOEs for use with the VICTORIA class combat system, and demonstrated the ability to leverage design based work for a secondary purpose. It is expected that the measures will be useful for a variety of other VICTORIA class support projects.

Future plans: The MOPS and MOEs developed under this contract will be used in the project experimentation to support investigations into C2 aboard the VICTORIA Class Submarine.

Sommaire

Development of Measures of Effectiveness and Performance from Cognitive Work Analysis Products

Gerald Lai; Tab Lamoureux; DRDC Atlantic CR 2011-282; R & D pour la défense Canada – Atlantique; février 2012.

Introduction : Recherche et développement pour la défense Canada (RDDC) – Atlantique est chargée d'examiner de nouveaux systèmes pour de commandement et contrôle (C2) à bord des sous-marins de la classe VICTORIA. Dans cette optique, RDDC Atlantique a fait appel à des techniques de facteurs humains telles que l'analyse du travail cognitif et l'analyse des tâches en fonction des objectifs, afin d'évaluer le travail effectué par l'équipe de commandement de la classe VICTORIA. Les travaux ont mené à la création d'un système intégré permettant d'afficher les renseignements de différentes sources d'information essentielle en même temps. RDDC Atlantique a également construit le simulateur virtuel « vVictoria » afin de tester la convivialité et l'utilité d'outils tels que le système intégré d'affichage de l'information (IIDS).

Le contrat actuel comportait deux objectifs : élaborer des mesures de l'efficacité (MOE) et du rendement (MOP) pour évaluer les activités C2 et déterminer si les analyses de travail réalisées pour la phase de conception peuvent servir à développer des MOE et MOP.

Résultats : Lors de la révision des résultats de l'analyse de travail pour la conception, 168 MOE et MOP ont été identifiées. Ces mesures ont été regroupées sous neuf domaines fonctionnels avant d'être vérifiées, examinées et améliorées par des sous-mariniers opérationnels. Les domaines fonctionnels étaient les suivants : communications (externes); connaissance de la situation (CS) commune (incluant l'interaction interne avec les membres de l'équipage); sécurité; manœuvre du navire; dissimulation; planification; gestion des contacts; CS individuelle; systèmes des sous-marins. Il a été conclu que les MOE, les MOP et les domaines fonctionnels étaient valides et qu'ils représentaient les activités effectuées à bord des sous-marins de la classe VICTORIA. Afin de catégoriser davantage les MOE et les MOP à la suite de cette validation, quatre scénarios de mission spécifique ont été créés : nagivation, gestion des contacts, planification (conjoncture favorable) et intervention en cas d'incident. Un sous-ensemble de mesures a été choisi pour chaque scénario afin d'optimiser la couverture des domaines fonctionnels. D'autres MOE et MOP ont été élaborées pour le IIDS dans les quatre scénarios.

Importance : La sélection de mesures informatives robustes est un élément essentiel de toutes les phases d'élaboration et d'essais. Un grand nombre de MOE et MOP potentielles pouvant être utilisées avec le système de combat de la classe VICTORIA a pris forme au cours de ce projet. Cela a également permis de démonter que le travail de conception déjà effectué est réutilisable. Les mesures devraient être utiles dans le cadre de divers autres projets de soutien de la classe VICTORIA.

Plans futurs : Les MOE et MOP élaborées durant ce projet serviront lors d'essais pour faciliter l'examen du C2 à bord des sous-marins de la classe VICTORIA.

Table of contents

List of figures

List of tables

This page intentionally left blank.

1 Introduction

1.1 Overview

Canadian Forces (CF) purchased four conventional diesel-electric UPHOLDER Class submarines from the Royal Navy in 1998 to replace the decommissioned OBERON Class of submarines. The new submarines were renamed VICTORIA Class and were delivered to Canada from December 2000 onwards (see Figure 1).

Figure 1: HMCS Windsor leaving Faslane, Scotland (ReadyAyeReady.com).

As with all military platforms, the expected operational life of the VICTORIA Class is long (in the region of 30 years or longer). This operational life can only be achieved through a diligently followed program of preventative and corrective maintenance. However, to maintain the submarine's operational effectiveness upgrades to sensor, weapons, and propulsion systems can be expected. In anticipation of these upgrades, Defence Research and Development Canada (DRDC) – Atlantic is conducting Research and Development (R&D) on the Command and Control (C2) systems in the submarine.

R&D on C2 is often focused on the digital information displays and controls provided to operators. This is important work because, with the advent of more capable and flexible sensor and weapons systems, there is more information that must be represented to the operator. It is desirable to present this information in integrated, composite representations that combine different data in ways that mimic an expert's internal mental model to support situation

awareness, problem solving, and decision making. This approach to information display reduces the mental workload, minimizes the chances of human error, and alleviates the display management activities that an operator must perform.

Since many of the sensor and weapons systems aboard vessels come from a variety of vendors, such integrated, composite displays are rarely provided. Rather, a systems integrator must build displays that integrate the different sources of data and present them coherently. Unfortunately, there are no rules concerning how this should be done. Therefore, R&D is necessary to support systems integration.

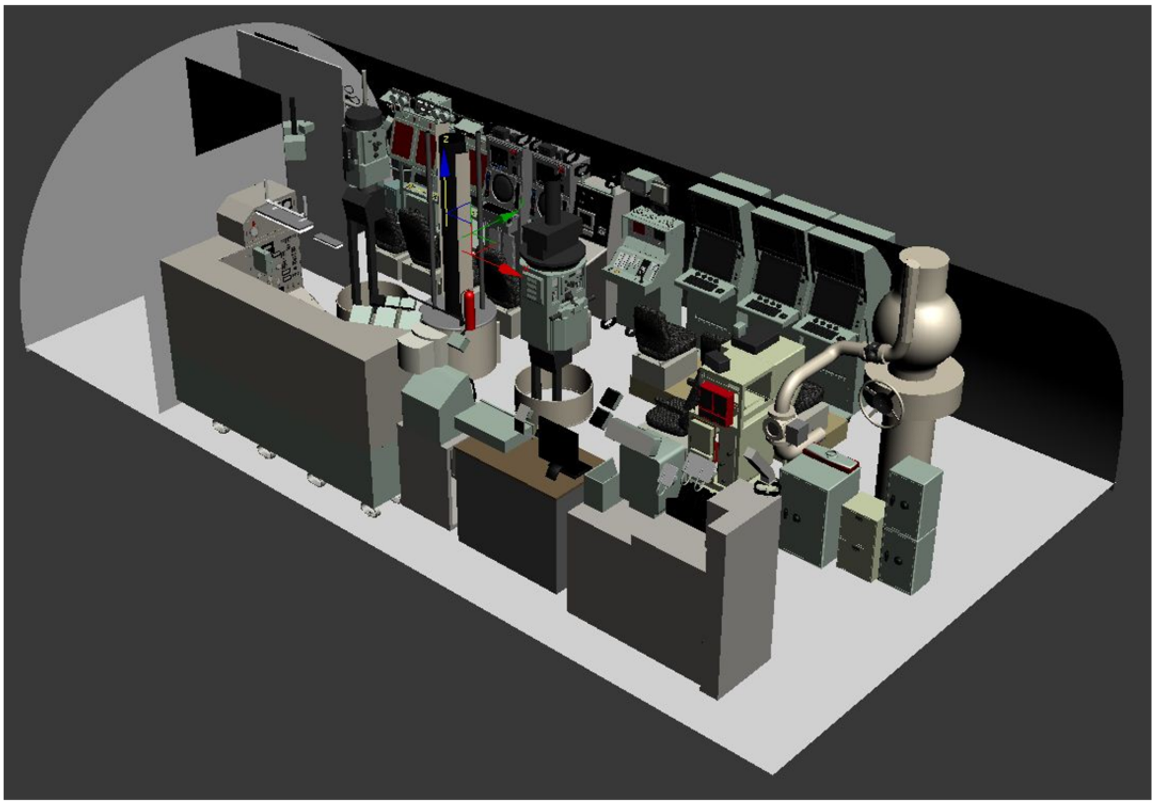

Figure 2: Three Dimensional Model of VICTORIA Class Control Room.

Introduction of new systems also changes the dynamics of the work aboard a submarine. For instance, integrated systems may change the manning/skills requirement, new equipment may affect the sightlines and communication, and operator tasks may evolve. To study new concepts in C2 systems, and understand the impact this may have on the team in a submarine control room, DRDC Atlantic is engaged in a number of efforts, including a series of experiments to be conducted in a full-size replica of the control room (see Figure 2 for a representation) being constructed at DRDC Atlantic. This simulation facility will faithfully replicate the systems, equipment, manning, and volume of space in the submarine control room, allowing DRDC Atlantic to effectively support the CF in extending the effective operational life of the VICTORIA Class submarine. The simulation facility is called Virtual Victoria (vVictoria or vVic).

DRDC Atlantic has previously carried out Cognitive Work Analysis (CWA) and Goal Directed Task Analysis (GDTA) of the activities in the VICTORIA control room. This work was carried with the intention of developing display and decision support systems that would improve the Situation Awareness (SA), problem solving and decision making of the control room team. It is DRDC Atlantic's plan to build prototype displays based on the concepts developed during this work and perform experiments to establish the benefits and shortcomings. To do this, Measures of Performance (MOPs) and Measures of Effectiveness (MOEs) must be defined and addressed during experimentation. MOPs and MOEs can be defined as follows:

- Measure of Effectiveness: A measure that characterizes the operational effectiveness of a unit or force in achieving its objectives during a mission. The measure must relate directly to the mission objectives and it must provide insight into the degree to which these objectives were satisfied. MOEs reflect a more holistic or overall measure of system performance in the achievement of its goal.
- Measure of Performance: This measures the performance of a particular system, and as such it is indirectly related to the mission objectives. It is usually related to technical properties of the analyzed systems, and should be consistent for corresponding systems across options. MOPs represent task performance at a lower level; that is, performance by individuals or small teams on activities that contribute to overall system success.

These definitions have been used throughout this project to guide the development of MOPs and MOEs.

1.2 Objectives of this Work

DRDC Atlantic contracted CAE Professional Services (CAE PS) to assist in the development of MOPs and MOEs for the vVic experiments. The objectives of this work are as follows:

a. Develop MOPs and MOEs for use in the vVic experimental facility for measurement of VICTORIA Class command team and control room performance; and,

b. Evaluate the efficacy of CWA and GDTA, performed for the purposes of design, as inputs to the development of MOPs and MOEs.

Elaborating on these requirements, the work required review of a variety of CWA and GDTA studies performed for DRDC Atlantic, the development of the MOPs and MOEs, and validation of the MOPs and MOEs with Subject Matter Experts (SMEs). More detail on the method is provided in Section 3.

1.3 This Document

This document is comprised of the following sections:

- Section 2 Previous CWA and GDTA Studies : this section describes the CWA and GDTA work leveraged in the development of MOPs and MOEs.
- Section 3 Method: this section describes the approach taken to develop MOPs and MOEs.

- Section 4 Results: this section organizes and describes the MOPs and MOEs developed for this project and presents an approach to running vVic experiments that facilitates the stimulation and measurement of the activities of interest.
- Section 5 Conclusions: discusses the findings and makes recommendations for the application of MOPs and MOEs to experiments and simulations in vVictoria.

2 Previous CWA and GDTA Studies

A significant amount of work concerned with understanding the activities in a VICTORIA Class control room from an operator perspective has already been completed. This work has been documented in five reports which trace:

- Two different analysis approaches (GDTA and CWA);
- The distillation of these analyses into critical work functions and requirements;
- The development of these requirements into display concepts;
- The selection of one design concept for development;
- The development of that design concept; and,
- The validation by SMEs of the analyses, the requirements, the competing design concepts, and the final design.

Below, each specific document is summarized with a particular emphasis placed on the description of what it provided to the development of MOPs and MOEs.

2.1 Analysis and Assessment of As-Is C3 for the VICTORIA Class Submarine (Taylor, Karthaus, & Bruyn Martin, 2009)

This report described the results of a GDTA study of the VICTORIA Class control room. GDTA is a method designed for the identification of SA requirements and information sources important to decision making and goal achievement (Endsley, Bolte & Jones, 2003). This study also carried out a Communications Flow Analysis. Insofar as the work described decisions, SA requirements and information sources, it did not describe the components of performance beyond the identification of a number of high-level functions. To facilitate measurement, a ‘component of performance’ would need to be an observable or measureable feature. The parameters associated with performance (good or bad) would be a specific quantity, value, or similar data that would be compared against the observed value to allow an evaluation to be made.

These functions were used to organize goals, decisions, SA requirements, and information sources. Nevertheless, the functions were useful in guiding the development of MOPs and MOEs. Specifically, it was expected that MOPs and MOEs would be developed for each function.

2.2 Identification of C3 Design Concepts for the VICTORIA Class Submarine (Bruyn Martin, Taylor, & Karthaus, 2009)

This report described the results of a CWA study of the VICTORIA Class control room. CWA is a method designed to identify the constraints that exist in a work domain and the opportunities for design that exist in the work practices adopted by skilled operators (Vicente, 1999). CWA is not founded on the existing work system (beyond the work practices of experts) but is meant to facilitate the development of truly innovative new systems, rather than the evolution of existing

systems. CWA consists of five inter-related analyses, of which three were completed for this work: the Work Domain Analysis (specifically a Contextual Activity Matrix analysis); Decision Ladder analysis (specifically a Cognitive Transformations Analysis); and a Strategies analysis. The resulting data was considered from both top-down and bottom-up perspectives in order to identify design concepts. These design concepts were then grouped according to their compatibility and elaborated for presentation to SMEs.

From the perspective of MOP and MOE development, this work reached a similar level to that of the GDTA: general work functions performed in the control room but little information about specific tasks. Again, this meant that the analysis fell short of identifying the measureable components of performance, and provided no information regarding the parameters of good and poor performance. The functions, however, augmented the list of functions derived from the GDTA and served as a verification of those functions that were found through both analysis approaches. There was not a perfect mapping.

2.3 Generation of Design Requirements for a C2 Information Integration Display (Bruyn Martin, Taylor, Karthaus, & Matthews, 2010)

This report elaborated of on the CWA work in order to understand the tasks performed in support of the work functions to the level required for design purposes. The specific work functions subject to elaboration were selected on the basis of the design concepts preferred by SMEs in the Bruyn Martin et al (2009) report. This work, however, focused on the information sources contributing to each of the functions supported by the preferred design concepts. This was required to progress the design. Although this level of detail is much lower than the functions described above, it still did not describe the components of performance nor the parameters of performance. The description of the information sources did, however, identify some likely sources of performance data, even if the description did not explain what performance would be measured or what an expected value might be.

2.4 Validation and Prioritization of C3 Concepts for the VICTORIA Class Submarine: Final Report (Bruyn Martin & Taylor, 2010)

This report describes the sessions held with SMEs to prioritize the design concepts identified in previous work. This was necessary in order to appropriately allocate effort for further development. This report provided very little additional information for the development of MOPs and MOEs.

2.5 Conceptual Design for the C2 Information Integration Display (Rehak, Karthaus, Lee, Matthews, & Taylor, 2011a)

This report described the Information Integration Display (IID) in detail, and included suggestions for experimentation. Many of these suggestions concerned decisions about how to represent information or where to locate information on the IID. These investigations were less

concerned with performance and effectiveness and, while interesting, are not relevant to the current development of MOPs and MOEs. However, the descriptions of the IID, when mapped to the functions selected for performance and effectiveness measurement in the current contract, provided a number of potential MOPs and MOEs. The number of potential MOPs and MOEs was also provided to DRDC Atlantic as information that could support decisions about which parts of the IID to build first.

2.6 Interaction Design and Storyboard Development for a C2 Information Integration Display: Final Report (Rehak, Karthaus, Lee, Matthews, & Taylor, 2011b)

This report mapped the CWA and GDTA analyses onto the design of IID. This display was the fully-realized concept selected previously by SMEs and developed by the authors. As such, this report provided very little additional information of value to the development of MOPs and MOEs.

2.7 Other References

Because the objective of this work was to develop MOPs and MOEs, it would not be satisfactory to depend solely upon the CWA and GDTA work once it became apparent that the information therein was insufficient to achieve the objective. Accordingly, additional VICTORIA Class analyses were consulted, notably Lamoureux, Pronovost, & Dubreuil (2011), as well as informal sources of information regarding VICTORIA Class control room activities.

2.8 Integration of Previous Work

Taken together, the previous CWA and GDTA work provided a broad perspective on the work activities of the command team aboard a VICTORIA Class submarine. Practically, however, the first report (Taylor, Karthaus, & Bruyn Martin, 2009) facilitated the distinction between MOPs and MOEs through the structure of the GDTA where goals, sub-goals, and functions corresponded to MOEs, and sub-sub-goals, decisions, and SA requirements corresponded to MOPs. The second report (Bruyn Martin, Taylor, & Karthaus, 2009) similarly facilitated the distinction on the basis of the CWA Contextual Activity Matrix. This analysis (subsequently extended for design purposes: Bruyn Martin, Taylor, Karthaus, & Matthews, 2010) focused on functions, tasks, and strategies. A composite of both the results of the GDTA and the CWA, and the most likely mapping of MOPs and MOEs to those composite results is presented in Figure 3.

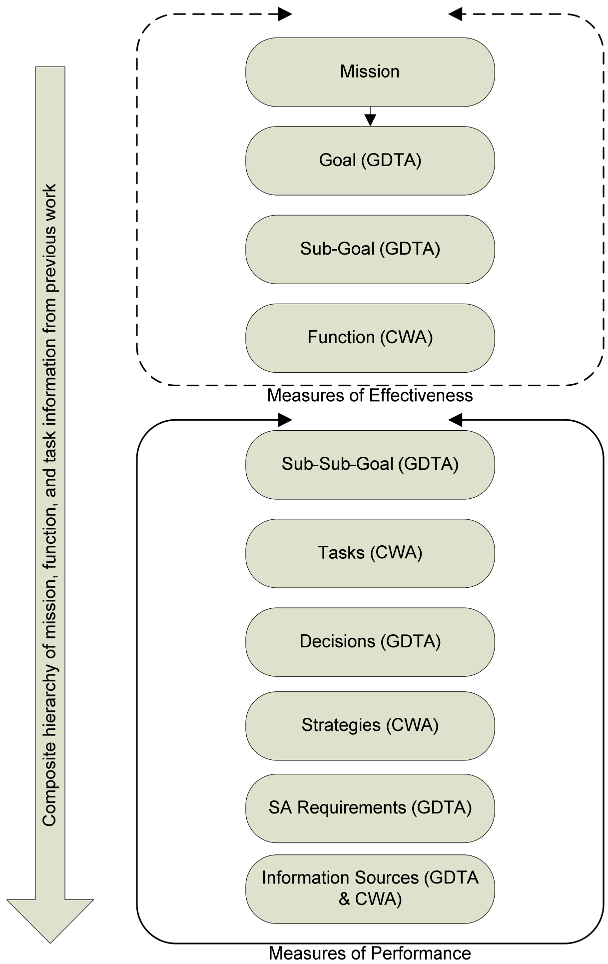

Figure 3: Relationship of Data in Previous CWA and GDTA Work.

GDTA outlines the following functions:

- Watch turnover;
- Blind pilotage;
- Command team management;
- Communications with TG or SUBOPATH;
- Detect and classify contact – dived;
- Detect and classify contact – surfaced or PD;
- Gathering and reporting intel;
- Incident detection and classification; and,
- Incident management.

CWA outlines the following mission phases:

- Planning;
- Transition;
- Execution;
- Post-execution; and,
- Return.

The mission phases could probably be applied to each of the functions.

CWA outlines the following functions:

- Ongoing mission planning and evaluation;
- Sensemaking and planning management;
- Command team management;
- Tactical and intelligence information acquisition;
- Contact management;
- Target identification and prioritization;
- Incident management;
- Target prosecution;
- Threat evasion and avoidance;
- Management of ownship tactical disposition (i.e., navigation);
- Ownship signature management;
- Facilitation of special forces operations;
- Platform systems management;
- External communications: tactical and operational information exchange; and,
- External communications: intelligence information exchange.

There isn't necessarily a mapping between the GDTA and the CWA. However, the CWA has the most information regarding tasks associated with the functions (Annex D of Bruyn Martin et al, 2009). Annex F (Bruyn Martin et al, 2009) also has a great deal of decision making information for each of the highest priority functions, but not related to performance or effectiveness. Annex G (Bruyn Martin et al, 2009) also describes strategies which could, with more focused work, be the genesis of additional MOPs and MOEs.

The Bruyn Martin et al (2010) report has some more detail regarding command activities:

- Track a contact of interest (COI);
- Predict future contact movement, behaviour, and intent;
- Integrate information related to tactical picture (building tactical SA);

- Monitor navigation with respect to the navigation plan;
- What is the ongoing tactical plan (concentration on snorting plan); and,
- Determine impact of submarine system degradation on mission goals.

The report (Bruyn Martin, 2010) also added a work function: overall tactical picture interpretation.

Extrapolating beyond the information provided by the previous work, there was a potential to develop more MOPs and MOEs than could be developed under a single, short-term contract. Thus, a method for prioritizing tasks, functions, and missions needed to be developed. Practically, MOPs and MOEs were developed for all functions and all goals. Thereafter, MOPs and MOEs were developed for selected sub-goals, sub-sub-goals, decisions, and SA requirements. These were prioritized according to the number of links each had on average across all the function hierarchies.

As well as measuring performance and effectiveness of the submarine command team, vVic will be used to perform experiments on the subject of the IID. Many of the MOPs and MOEs developed for the functions and tasks described above should be sufficiently sensitive to identify differences that can be attributed to the application of the IID. These would likely focus on output measures of quality for tasks that would involve traditional information sources which have been presented using the IID. However, should the interest be in the efficacy of the design of the IID, MOPs and MOEs would need to focus specifically on the IID. MOPs for the IID can focus on Eye Movement Tracking (EMT), mouse movements and ‘dwells’ (of either the mouse or gaze), as well as comparing performance data for specific tasks between control and experimental conditions.

3 Method

The development of MOPs and MOEs for vVic began with the review of previous work, the results of which are presented in Section 2. While some additional knowledge was brought from previous work conducted by CAE PS, no additional sources of potential MOPs and MOEs (in particular CF training documents and evaluation guidance) were used in this study. The information gathered from the previous work was considered for what it contributed directly or indirectly to MOPs and MOEs development. This consideration followed the process described in Section 3.1.

It was felt that MOPs and MOEs should be developed for the categories found in the previous work (Section 3.2). Additionally, it was felt that both subjective and objective measures should be developed. Once the creation of MOPs and MOEs was complete, and any additional thoughts or concerns were captured, the MOPs and MOEs were compiled in a package suitable for validation with SMEs (Section 3.3). Accordingly, SME validation interviews were held over two days in Halifax and Dartmouth. The SME validation involved individual interviews with SMEs and a final group debrief in which findings were presented and further comments and discussions were invited.

Based on the SME validation sessions and the analysis work done to create the MOPs and MOEs, further analysis was done to settle upon an approach that would facilitate MOP and MOE collection (Section 4). Because the SME feedback received was predominantly qualitative, analysis focused on trends and significant issues noted by the SMEs and how to integrate these features with the requirements of MOP and MOE collection into a coherent measurement approach (Section 4).

3.1 Process Flowchart

Given it was not clear at the outset what information would be in the previous CWA and GDTA work the process outlined on the following page evolved to guide the project. Had clear standards for performance or task sequences been described, MOPs and MOEs would have been straightforward to develop. This was not the case. Therefore, an objective of this study became to develop a systematic approach to the development of MOPs and MOEs, based on the analysis of the mission, functions and tasks already performed. The systematic approach was developed as a flow chart, depicted in Figure 4. This approach is based on a consideration of the contractors' own thought processes, after a search of the literature for an existing approach to MOP and MOE development was unsuccessful.

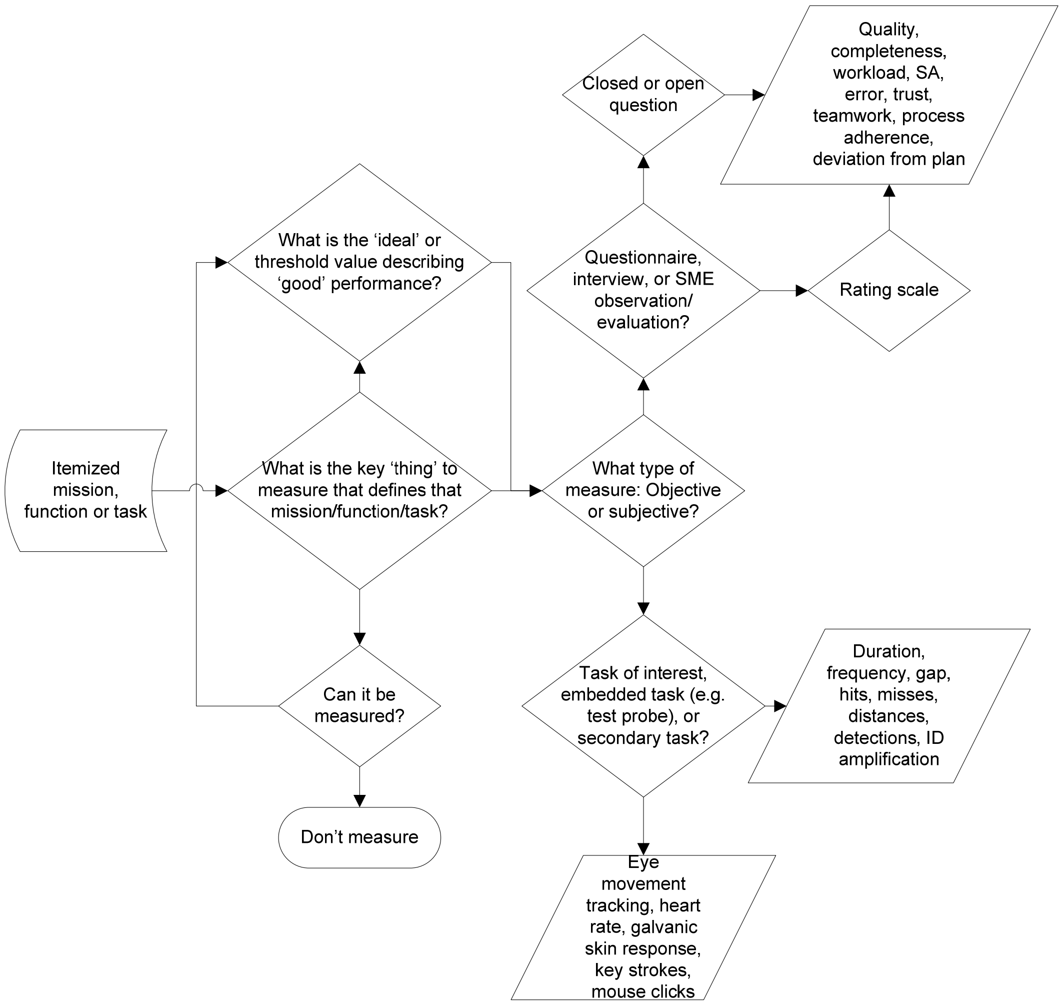

Figure 4: Flowchart for MOP/MOE Development.

The flowchart attempts to list all the possible types of measures that might be used to assess command team performance in the VICTORIA Class control room. Note that this is not an exhaustive list of possible measures; rather it is an extensive list of *types* of measures. The specific measures to be used necessarily take account of the scenario, the mission, and the likely tasks to be performed. Figure 4 also makes this clear with the decision to be made about the key 'thing' that defines that mission/function/task.

This schematic is useful to illustrate the development of MOPs and MOEs and to ensure that each MOP and MOE was arrived at systematically.

3.2 Categorical Organization

The abstraction of goals and functions reported in the previous work did not lead to obvious metrics of performance and effectiveness. This was due to different analysis methods, as well as varying levels of categorization, ranging from broad to specific. The content within these reports were nevertheless quite useful in informing the development of MOPs and MOEs.

For the purposes of developing MOPs and MOEs, as well as presenting MOPs and MOEs to SMEs for validation, it was necessary to capture existing goals or functions relevant to MOPs/MOEs and to construct a meaningful framework under which to group areas of work. The intent was not to critique the 'labels' used to organize, allocate, or describe them, but rather to establish and identify commonalities;

The approach was to compile all relevant goal or function groupings that could conceivably be mapped to an observable form of performance or effectiveness, regardless of the goal level (sub-goal, sub-sub goals, function, etc.). This data exists as itemized line entries, sorted within an Excel spreadsheet.

Nine resulting groupings were developed to guide MOP and MOE development, and to structure the presentation to SMEs for validation. These work areas are:

- Communications (external);
- Shared SA (includes internal crew interaction throughout the submarine);
- Safety;
- Seamanship (subsequently Ship Handling, based on SME feedback);
- Covertness;
- Planning;
- Contact Management;
- Individual SA; and,
- Submarine Systems.

3.3 SME Interviews

Interviews were conducted over two days at DRDC Atlantic and at the Naval Dockyard in Halifax. A total of four SMEs were involved, all with command experience aboard the VICTORIA Class. SMEs were interviewed individually, and then brought together for a group session at the conclusion of the individual interviews.

3.3.1 Mission Template

A general mission overview was presented to the SMEs to allow them to consider the MOPs and MOEs in context. The mission description is as follows:

- Covert surveillance and tracking scenario;
- Submarine has been tasked to conduct covert surveillance and tracking of maritime traffic in the approaches to a non-allied neutral country;
- The mission is to obtain information on the normal activity in the area and identify any anomalous behaviour;
- The area of operations is a 25 x 40 nm box south of the port, other North Atlantic Treaty Organization (NATO) units will be operating in adjacent water areas;

- Unit will maintain normal communications cycles;
- Local units may or may not be operating Automatic Information Service (AIS) in accordance with International Maritime Organization (IMO) regulations and space-based radar contact maps are available twice daily;
- Unit is authorized to enter national waters if required but remaining covert is deemed essential;
- National maritime air assets are shore and ship launched helicopter. No known Maritime Patrol Aircraft (MPA) assets in the area. National naval units may be operating in the area;
- Water depths 50 - 150 m, shallow surface duct (15m). Weather is overcast, wind from south 15kts, SS 3-4. Forecast is to stay the same. Tides are 2-4 m, High Tide at 0900;
- Expect Heavy shipping in channel, Light shipping into port. Heavy local fishing. Light pleasure boat traffic;
- Conduct reconnaissance of a highly classified and sensitive target;
- Areas of shallow waters and confined spaces;
- May be exposed to mine area (possibly inconsistent with the rest of mission narrative); and,
- Torpedo engagements real possibility (possibly inconsistent with the rest of mission narrative).

Note that this scenario was presented during the validation sessions and is expected to change and evolve before being sufficient for vVic experiments. Each SME identified some inconsistencies with the scenario which will be valuable when developing an experimental scenario for vVic. The critiquing of the scenario also highlighted some potential MOPs that were subsequently developed. Anecdotally, this is often a useful secondary outcome of developing scenarios for presentation to SMEs.

3.3.2 Work Area Descriptions

A description of each work function was provided to the SMEs for their validation. This was necessary to ensure that we were attempting to develop MOPs and MOEs for an accurate and correctly bounded concept. After agreeing or modifying the description, the SMEs were presented with MOPs, split into objective and subjective, and then MOEs, also split into objective and subjective. The presentation used with the SMEs is included in Annex B.

3.3.3 Objective/Subjective Measures

There are two general ways to measure performance and effectiveness: objectively and subjectively. Objective measures are often collected automatically and are generally those qualities or quantities that can be captured, measured, or otherwise recorded by a secondary instrument. In other words, they do not rely on the self-report of the person being measured. Objective measures require a great deal of thought beforehand in order to set in place the correct

measurement instrument, use the correct unit of measurement, and identify how and when to take measurements. Objective measures are often unobtrusive and provide immediate answers to the question being asked.

Subjective measurements rely on the self-report or observation of the person being measured. This self-report can be on a numerical scale, binary (e.g., yes/no, true/false) answers, or can be completely open-ended, allowing the person being measured scope to qualify and caveat their answer. Subjective measurement is often more intrusive because the participant has to ignore the task for a period of time, or else they are retrospective because the measure is administered at the conclusion of an exercise. Subjective measures still require care and thought to set up, but the focus is on the adequacy of the questions, rather than on the mechanism of collection. Subjective measures often require significant analysis after collection in order to arrive at the answers to a question, but they are a potentially rich source of supplemental information.

Many data of interest in Human Factors studies can only be collected through subjective means. For instance, SA is different for all people, so it must be collected subjectively to allow the full range of possibilities. Likewise, investigations into cognition and decision making provide the most data when approached using subjective methods (while objective methods can be used to measure performance, understanding performance can only be achieved through subjective methods).

3.3.4 Rating Approach

Initially, the plan was to ask SMEs to rate the functions, the MOPs and the MOEs. Each measurement area would be described according to the defining characteristics of the function to be measured and the SME would rate their agreement or disagreement on a scale of 1 to 5, with any alternative descriptions noted. Then, for each measurement area, the MOPs and MOEs would be described and the SMEs would rate the 'appropriateness' of the MOP/MOE (also on a scale of 1 to 5), suggesting any alternative measurements they believe are equally or more suitable.

This approach to validating the MOPs and MOEs, as well as the understanding of the different functions, was determined to be unworkable given the amount of time available with each SME. As such, no SME was presented with all function descriptions, MOPs and MOEs; rather, each SME was presented with a selection of functions representing those that were a) considered most important or b) had not been considered as fully as the others. Obviously, this latter criteria became more likely with each subsequent SME. The final tally of how often each function was discussed is shown in Table 1.

Table 1: Frequency of Discussion of each Function.

Function	Discussed	Function	Discussed
Communications	2	Shared SA	3
Safety	2	Ship Handling	2
Covertness	3	Planning	2
Contact Management	2	Individual SA	3
Submarine Systems	2		

Also, experience with the first SME interviewed indicated that sessions were likely to be more conversational than structured; otherwise it would not be possible to obtain good additional MOPs and MOEs. To have imposed the rating approach would have artificially ended each discussion about a particular function. Nevertheless, each SME was asked whether the description of the function was appropriate.

4 Results

This contract sought to achieve two objectives: to develop MOPs and MOEs, and to assess the ability to CWA work to support MOP and MOE development. This section discusses the findings for each objective, beginning with the assumptions, limitations and constraints that influenced the project, before moving on to a discussion of the adequacy of CWA in supporting MOP/MOE development, discussion of the analysis approach, and finishing with the measurement approach recommended by this contract.

4.1 Assumptions, Limitations, Constraints

The following list documents any assumptions, limitations, or constraints of the project during the execution of work and any related considerations for follow on work:

4.1.1 Assumptions

- The MOP/MOEs identified in this report are intended to support the development of experimental planning and execution within the vVic simulator located at DRDC-Atlantic;
- The MOP/MOEs presented in this report are valid measures that can be used for experimental planning;
- The template mission scenario presented during the interviews provided enough context for SMEs to provide insightful feedback on the proposed MOP/MOEs;
- During the interviews it was necessary to assume that external stimuli (outside of the submarine) could be represented to some degree of fidelity. This enabled SMEs to recommend scenario contexts that would be required to achieve many of the proposed MOP/MOEs;
- SMEs will be available in the future for consultation to develop scenarios at a level of detail needed to support experimentation. This ensures that the behaviours driven by scenario events are aligned with research objectives and subsequent measures of performance; and,
- Despite leveraging CWA work products; there was no intention for the current study to employ specific CWA methods to develop MOPs and MOEs.

4.1.2 Limitations

- It was not possible to illustrate examples of design interventions using prototype interfaces as these were in the early stages of development. Only general definitions pertaining to the interface capability areas were available at the time of this report;
- The list of proposed MOP/MOEs presented in this report are not intended to be an exhaustive list of measures for experimentation, but as a targeted selection of the most relevant and plausible areas for testing; and,

- No reference was made to training documentation or evaluation approaches to supplement the information provided by the previous CWA and GDTA work.

4.1.3 Constraints

- Four SMEs were interviewed, using a 2 hour time window for each participant. This timeframe allowed for the proposed MOP/MOEs to be discussed as groupings under the functions proposed, but did not allow for itemized inspection of each metric. Exhaustive discussion of each metric was deemed as unnecessary and time consuming;
- Psycho-physiological measures were deemed to be less feasible and outside the scope of future experimentation. These types of measures were not emphasized during discussion with the SMEs;
- The recommendations within this report should be used as a guideline for future experimental planning rather than to delineate experimental conditions. It was difficult to define test conditions, particularly at the design intervention level as the vVic interfaces were in their infancy and further understanding of their capability was required;
- A general outline of a typical mission profile was used to frame discussions during the interview sessions. Future experimentation will require customized scenarios to better isolate the desired metrics intended to reflect performance; and,
- The majority of proposed MOP/MOEs do not attempt to differentiate between performance evaluation at the team or individual level. This constraint acknowledges that future research objectives will need to dictate the level of complexity as it relates to team or individuals when assessing performance.

4.2 Use of Previous CWA and GDTA Work in the Development of MOPs and MOEs

The previous work (Taylor et al, 2009; Bruyn Martin et al, 2009; Bruyn Martin et al, 2010; Bruyn Martin & Taylor, 2010; Rehak et al, 2011a; Rehak et al, 2011b) was extremely detailed with respect to the manner in which it developed descriptions of the work domain. At the functional level of description, the analyses could suggest more MOP and MOE topics than could be developed under a single, short-term contract. The analyses were conducted for the purposes of design, however, and the level of detail was not sufficient for the development of measures beyond the simple identification of these functions and measurement topics. This section explains this finding in more detail.

4.2.1 Goal Directed Task Analysis

This analysis used SME interviews to identify those areas of activity that were deemed generically representative of the activities of the VICTORIA Class Submarine Control Room. The SME interviews took two approaches to collecting this information: by asking SMEs to identify these functional areas, and by asking SMEs to describe discrete elements on their role. The latter approach required the analysts to aggregate discrete elements into functional areas,

thereby validating what the SME provided when asked specifically to identify the functional areas. Much of the questioning resulted in convergent validation of the information being gathered by other means. The analysts also carried out a Communication Flow Analysis as a method of understanding the inter-relationships between ‘actors’ in the VICTORIA Class Submarine work domain.

The question set employed by the analysts for both approaches focused on describing the work domain with respect to its major components and the inter-relations between components. The Communication Flow Analysis further described these inter-relationships. The question set did not address the parameters of these components that defined good or bad performance. This is not to say that the analysis did not proceed to a significant level of detail. Indeed, for each functional area the analysts identified the goals, subgoals, decisions, SA requirements, and information sources. Rather, the nature of the description of each discrete element did not answer questions about performance and effectiveness.

4.2.1.1 Communication Flow Analysis

Consider a Communication Flow Analysis example. Figure 5 describes the various communications that may occur when surfaced, the medium by which they are made, and the participants in the communication.

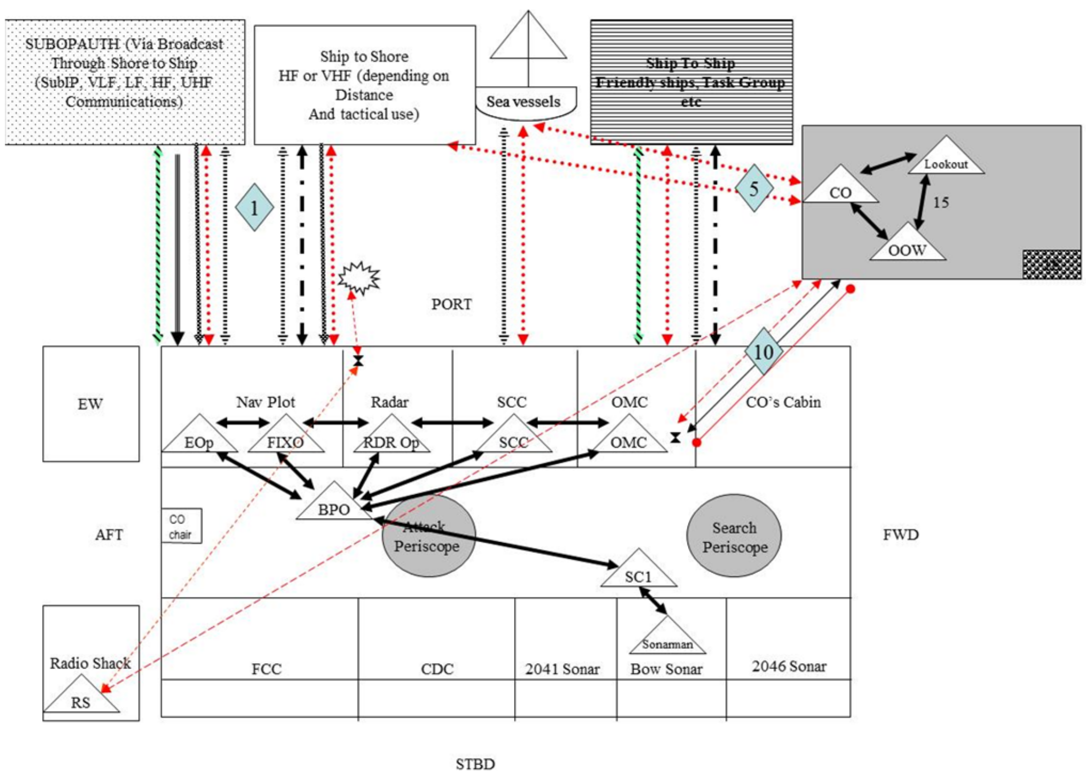

Figure 5: Surface Communications Aboard a VICTORIA Class Submarine (from Taylor, et al. 2009).

The text associated with the Communications Flow Analysis is concerned with further describing factors that may challenge communications. The analysis or description does not provide information concerned with how the submariner would determine communication was successful, unsuccessful, or in some other way sub-optimal. Technical challenges are listed but these are of limited value for simulation or experimental evaluation of human performance. Parameters such as duration, frequency (not radio frequency), content, and subsequent actions would be directly useful for developing MOPs and MOEs.

4.2.1.2 Goal Directed Task Analysis

Consider a GDTA example. Each GDTA analysis is extensive, which precludes the presentation of entire GDTA. Rather, a single 'path' through a GDTA is described in Table 2: GDTA Example. The number in brackets in the second column indicates how many total items have been described at that level in the analysis.

Table 2: GDTA Example.

Function	Detect and Classify Contact – Dived (9)
Goals	Detect and classify contacts (3/3 subgoals)
Decisions	Where is the new contact? (5/1 sub-decision)
Sub-Sub-Decision	Is there a new contact?
SA Requirements	Oceanographic conditions; Current bearing of contacts; Current course of contacts; Acoustic signature of contacts (12/7 sub-SA Requirements)
Information Sources	Bathy display (for Oceanographic conditions) (15 information sources)

Ignoring analysis oddities such as the decision concerning whether a contact has actually been detected being the single constituent decision of deciding where the contact is, the ensuring SA requirements do not say what specific information will lead to SA. For instance, what is the submariner looking for on the bathy display that will assist them to determine that there is or isn't a new contact? How is oceanographic information used with current bearing, course and acoustic signature to determine that a contact is actually real? While the listing of decisions allows good MOPs to be developed (i.e., whether a correct or incorrect decision is made), they will be highly dependent upon the scenario constructed and knowing ground truth. To directly assist the development of MOPs and MOEs that are applicable independent of detailed knowledge and control of the scenario, involves knowing what the specific components of SA are, what the values are, how swiftly the operator needs to make the decision, and what each information source provides to SA that allows the decision to be made. This list is not exhaustive. Further,

knowing this information would not only assist the development of MOPs and MOEs, it would assist the process of designing new interface concepts as well as ultimately building those concepts.

Ultimately, the GDTA identified the broad areas against which MOPs and MOEs should be developed. Specifically, the work functions identified, the goals, and some decisions. These could be used with the CWA output as described below.

4.2.2 Cognitive Work Analysis

This analysis again used SME interviews to gather information and involved a Work Organization Analysis (resulting in a Contextual Activity Matrix [CAM]), a Cognitive Transformations Analysis (CogTA), and a Strategies Analysis. The CogTA and the Strategies Analysis are related in that they describe "what needs to be done" and "how can it be done" respectively. Of particular interest, scenario-based semi-structured interviews were employed to guide discussion.

4.2.2.1 Contextual Activity Matrix

Initial work functions developed during the GDTA were used in the development of scenarios and question sets for use with SMEs. The CAM that resulted from this phase of work was not subject to any real analysis, although it did provide a list of work situations that is qualitatively orthogonal to the work functions. These work situations were planning phase, transition phase, execution phase, post-execution phase, and return phase. Although these categories do not directly develop MOPs or MOEs, they do help ensure that the MOPs and MOEs are developed to address all work situations, particularly when the work situations are overlaid with the work functions from the GDTA to create a matrix. This structure is a guide to analysts when verifying that MOPs and MOEs have achieved adequate coverage of the work domain.

4.2.2.2 Cognitive Transformations Analysis

The CogTA in many ways replicated the GDTA. That is, the GDTA listed goals, decision, SA requirements, and information sources, while the decision ladder in CogTA describes the operators' progress from detection, through development of SA, to decision and action, using different states of knowledge and information processing activities along the way (i.e., cognitive transformations). However, many of the states of knowledge and the information processing activities are described in the same terms as the decisions and SA requirements described in the GDTA. For instance, Table 3 is an excerpt from Table 9 of Bruyn Martin et al (2009).

Table 3: Excerpt of CogTA (Table 9 of Bruyn Martin et al, 2009).

Step Type	Decision Ladder Step	Details
Information-processing activity	Identify state	What is the situation? Is it confirmed as an emergency? What is the severity of the situation? What is the extent of the damage? What stage of incident management is required? (detecting, identifying, assessing, taking action) What is the current tactical situation? Are there threats? What are current tactical activities? (e.g., engaging a target) What is the internal damage? What are the priorities? What is the state of the platform systems?
State of knowledge	SYSTEM STATE	Alarms on the surveillance systems (e.g., flood alarms in bilges), \ crew members warning (e.g., yelling, informative pipe) machinery operation outside safe parameter, flooding, fire, smells outside normal parameters (e.g., electrical burning)
Information-processing activity	Predict consequences	Positive: Emergency is properly identified and assessed. Tactical situation allows desired emergency response. Emergency actions successfully control and/or rectify the internal emergency. Negative: The emergency is not properly identified or assessed and/or actions do not successfully control or rectify the internal emergency. Tactical situation, state of submarine, or state of crew does not allow desired emergency response.
State of knowledge	OPTIONS	If the incident is real (i.e., not a false alarm), further action must be taken if situation warrants such a response. If it is a false alarm, no further action must be taken. If it is an emergency can be contained and rectified quickly, HQ may not be stood up. If emergency cannot be contained and rectified quickly, First Headquarters (HQ1) should be stood up. Consider tradeoffs between safety vs. mission vs. covertness. If incident jeopardizes safety of crew, safety will override goal of achieving mission or covertness (e.g., must surface if deep or at PD). If crew safety is not at risk, CO will decide whether incident can be effectively managed while still accomplishing mission and/or remaining covert.

None of the details in the third column provide enough information to develop an MOP or MOE that is independent of the specific context of the scenario. As with the GDTA, the correct answer is highly dependent on the scenario, the experimenter's control of the scenario, and their knowledge of ground truth and the correct answer. However, the CWA approach provides a strategies analysis detailing how a task can be carried out.

4.2.2.3 Strategies Analysis

Strategies are introduced as possible ‘leaps’ and ‘shunts’ (shortcuts in the decision ladder) before being the subject of a more detailed analysis. An excerpt of the leaps and shunts table (Table 10 in Bruyn Martin et al, 2009) is presented in Table 4.

Table 4: Excerpt of Leaps and Shunts (Table 9 of Bruyn Martin et al, 2009).

Work Function	Leaps	Shunts
Ongoing mission planning and evaluation	**Leap: System State to Procedure** When adequate conditions meet the requirements for action in pre-established Standing Operating Procedures (SOPs), EOPs, or other relevant procedures may be carried out immediately	**Shunt: Evaluate Performance to Information** When no-go decisions are chosen due to inadequate conditions other routes may be chosen though may require progression through Information up through Evaluate Performance. **Shunt: Evaluate Performance to Procedures** When no-go decisions are chosen due to adequate conditions, relevant SOPs, EOPs and other procedures may be carried out immediately.

Again, analyzing the description leads to questions: What are “adequate conditions”? What are “relevant procedures”? What counts as ‘immediately”? What are the “no-go decisions” to be made? Should the operator routinely have “other routes” in mind? How many? How quickly should the operator decide to carry out “relevant SOPs, EOPs, and other procedures”?

The full strategies analysis was divided into General Strategies and Specific Strategies. Both types were described in a similar manner: the purpose of the strategy, the strategy itself, and the contextual factors affecting strategy choice. Table 5 and Table 6 present excerpts of tables describing general strategies and specific strategies.

Table 5: Excerpt of General Strategies (Annex G of Bruyn Martin et al, 2009).

Purpose of Strategies	Strategies	Contextual Factors Affecting Strategy Choice	Potential Design Concepts and Constraints
Predict outcomes	Consult memory, team consult, seek additional information about task, seek additional information about tactical situation	Own experience, experience of team, trust in team, seriousness of task/situation, tactical situation, typicality of situation, uncertainty of information, platform state	Outcome prediction will have to be supported in formats that will be able to be used by multiple users to enable team consult approaches

The information provided in the strategies analysis is very detailed and serves as good guide for design, particularly the contextual factors. Additionally, the analysis points to where MOPs and MOEs would be valuable, but it does not provide much insight into the strategies an expert might adopt (i.e., mnemonics and similar ‘tricks’ to facilitate good performance). Rather, they seem to represent a finer level of task detail than is provided elsewhere in the reports. The listing of tasks and contextual factors does not provide any of the detail required to develop MOPs or MOEs that

can be applied independent of the scenario. Again, the experimenter would need to maintain tight control over the scenario and know what ground truth is when any measurement is taken to understand the impact on performance.

The listing of contextual factors that can affect strategies does, however, assist in the development of MOPs and MOEs that are predicated on the scenario. The contextual factors can be used when constructing a scenario to challenge the operator. By measuring a gross level of performance (e.g., time to carry out an activity, where the onset and completion of the activity are easily identified) under experimental conditions in which the contextual factors are varied, MOPs and MOEs can shed light on issues that will have an impact on system design, training design, procedures, and team work.

Table 6: Excerpt of Specific Strategies (Interpret Overall Tactical Picture, Annex G of Bruyn Martin et al, 2009).

Decision Ladder Step Type	Decision Ladder Step	Interpret Overall Tactical Picture	Purpose of Strategies	Strategies	Contextual Factors Affecting Strategy Choice	Potential Design Concepts and Constraints
Information processing activity	**Predict Consequences**	Positive: Overall tactical picture is interpreted correctly and more information is gathered or actions are taken to maintain safety of boat and crew and achieve mission. Negative: Overall tactical picture is interpreted incorrectly and safety of boat and/or crew may be at risk or mission may not be achieved. Additional information that is needed is not gathered.	Create timeline for updating tactical picture	Personal preference, procedures	Time pressure, availability of procedures, tactical context	
			Compare information presented in different ways on different displays	Prioritize one display and ignore others; prioritize representations equally and integrate in head, integrate all into one display, team consult	Time pressure, availability of technologies, procedures, trust in team, personal preference, trust in own judgment	Need consistency between displays to extent possible - same symbology
			Update tactical picture	Updating navigation chart, updating fire control, manual updates, automatic updates, updating tactical picture on dedicated display, updating tactical picture on plans display	Available technology, available information, procedures, time pressure	Technologies for representing integrated tactical picture and integrating with plans

4.2.3 Leveraging GDTA and CWA for MOP/MOE Development

As noted throughout this document, the previous GDTA and CWA work was useful in identifying a framework within which to develop MOPs and MOEs (i.e., the functional areas and the work situations). Using this framework as a guide, coupled with pre-existing knowledge regarding MOPs, MOEs, objective and subjective measures, and typical qualities of measurement interest (e.g., SA, workload, errors, accuracy, time), it is relatively easy to develop specific ideas for MOPs and MOEs. However, the previous work did not provide a further level of detail defining

successful, unsuccessful, or otherwise sub-optimal performance, meaning that the previous work could not be used directly as a source of 'absolute' MOPs and MOEs (i.e., measures that clearly indicate good or poor performance, success or failure, irrespective of experimental condition). Further, it was difficult to draw conclusions from the previous work that would allow the development of situationally-specific MOPs and MOEs. The analysis that came closest to achieve this was the Strategies Analysis, particularly when conducted against the CogTA.

The previous analyses could be extremely useful, however, if additional opportunities were available to work with SMEs or review training, evaluation, and performance data. Having used the previous analyses to narrow the search for MOPs and MOEs, an analyst team could now develop a very specific agenda to guide the further development of MOPs and MOEs. This was not the approach, nor were there opportunities, in the current contract. Of particular use would be the opportunity to observe instructors carry out evaluations of training, and to conduct interviews with them after the evaluation to access the 'meta-knowledge' by which they make their evaluation. This would obviously be additional effort but, if carried out by members of the original analysis team, much of the familiarization required for this contract would be obviated and the net additional cost would probably be equal or less to the value of the current contract.

4.3 Post Interview Analysis

The MOPs and MOEs that have been developed are intended to represent as broad and deep coverage as possible within the scope of this contract. In this respect, each function (i.e., measurement area) was defined to be as independent of the other functions as possible. As a consequence, many MOPs and MOEs were possible, but a subset had to be selected for development. This section describes how this subset was selected, before the next section describes a method for applying the measures and the measures themselves.

4.3.1 SME Interviews

The SME interview session facilitated the validation of proposed MOPs and MOEs, and provided further insight regarding potential methods of quantifying the different measurement areas. Four officers were interviewed: three Lieutenant-Commanders and a Commander. Three had CF command experience aboard submarines; one was currently an Executive Officer (XO) aboard a submarine. One officer is the current Submarine Operating Authority (SUBOPAUTH) and another is currently responsible for submarine tactics at the Maritime Warfare Centre in Halifax. A DRDC Human Research Ethics protocol for the interviews was submitted and reviewed. (Hunter 2011).

As noted in Table 1, all functions were discussed with two or three SMEs. Although each interview began with a defined agenda, the actual process was more conversational and followed the flow of information provided by the SME. This meant that functions were addressed in a different order for each SME, depending upon the manner in which they related the topics. The interviews were actively directed only to the extent required to address each function an equal number of times. Between 3 and 7 functions were discussed in each interview for a total of 21 discussions of functions.

There was general consensus among the SMEs that the categories presented during interviews provided sufficient coverage and representation of the work areas expected of the Commanding

Officer (CO) or Officer Of the Watch (OOW) on the VICTORIA Class Submarine. All participants expressed general agreement with the measures proposed, with the exception of select terminology which were updated on the presentation slides during the sessions. A number of key points were made during the interviews. These points guided the subsequent development of MOPs and MOEs.

- Safety and covertness are the overarching and guiding principles that influence how all other constituent work tasks are executed. One participant summarized the priority of submarine goals as follows: "The safety of the submarine and her personnel, remain undetected, and achieve the aim, in that order". The CO will often give an order and phrase it in the context of this hierarchy in order to communicate with the team where the priority is. The goal priorities above may vary according to the CO's assessment of the situation and environment.
- Planning and ship handling are both important functions, particularly with respect to maintaining and re-gaining tactical advantage in the event of a counter-detection. However, there is rarely a 'right' answer to how a mission should be planned or how the boat should be handled. For instance, a deviation from course is often acceptable when traded off against the boat's tactical disposition such as when the boat is turned as part of the stern arcs clearance procedure or to utilize features of the environment (e.g., land, sea mounts) to further minimize chance of detection.
- The conditions under which measurements will be taken will not include the impact on vigilance and fatigue that long periods at sea and on watch have. These will be difficult to achieve in experiments but must be acknowledged as key influences on performance.
- The 'elegance'[1] of planning can be assessed by the number of concurrent activities the submarine is able to do at once or by the amount of information it can collect in a given time window. For instance, during a single surfacing, being able to empty the bilges, conduct communications, snort, conduct an all-sensor search, collect electronic warfare information, update the weather picture (wind and rain), and project how this might affect the boat. The development of tools that facilitate the speed at which activities can be completed could provide interesting opportunities for comparative studies.
- Situation Reports (Sit Reps) are streams of consciousness that are often rich sources of information but are difficult to assess. Sit Reps may provide experimenters insight into the appropriate level of communications required to achieve optimal SA among the crew.
- Given that submarines operate in an 'open' environment it is very difficult to pre-define the 'ideal' answer to a subjective query. By controlling the experimental scenario the investigator is effectively creating a 'closed' environment and can therefore know ground truth. Thus, 'ideal' answers to subjective queries can be pre-defined.
- Real-time probe questions asked by the investigator provide insight into the information processing activities of the operator and may not necessarily be intrusive given an operator's training to focus on the primary task (i.e., to ignore low priority interruptions). However, it will be difficult to maintain inter- and intra-rater reliability with such subjective measures.

---

[1] Elegance is the attribute of being unusually effective and simple. Commonly used in mathematics, engineering, and psychology, in the context of submarine planning it refers to the ability to achieve many goals contemporaneously.

- The scenario profile needs to be tailored for the MOP and/or MOE under consideration. Specifically, the scenario will need to include event-driven aspects that necessitate the subject's attention to specific task components. For instance, covertness is not always critical when no threat will result from being detected or when the objective is to actually be detected (for safety or 'show of force' reasons).
- There is a general need for operators aboard a submarine to consider the fire control solution as well as the TMA solution. Specifically, the fire control solution helps to validate the TMA solution. This cross-check emphasizes the importance of mental arithmetic abilities which can help to expedite solutions when required (for instance, when the submarine needs to return to periscope depth quickly). The need also implies the possibility of complacency and reliance on computed solutions throughout the boat. Submariners need a better fundamental understanding of the key pieces of information presented to them, not just an ability to 'stack the dots' or similar decision-support activities.
- SA for submarine systems is significantly enhanced by having a good schematic representation. Such a schematic is particularly useful for translating any faults into an understanding of what the boat remains capable of doing and what it can no longer do. This is not part of the IID but may be desirable in the future and would require particular requirements of an experiment to support evaluation.

The SMEs had no significant objection to any of the functional areas identified but were less able to provide unequivocal validation of the MOPs and MOEs. The MOPs and MOEs will need to be assessed for their feasibility in the vVic simulator and their likely value in discriminating different experimental conditions.

The post interview analysis involved further reduction of the feedback gathered from the SMEs and structured with experimental planning in mind. Considerations pertaining to operational research (user experimentation in vVic simulator) are used to develop a framework of analysis and reporting. These considerations are discussed below.

4.3.2 Feasibility Evaluation of MOPs and MOEs

Determination of MOP/MOE suitability ratings were not obtained from SMEs during the interviews due to constraints related to time and prioritization of feedback. It however remained necessary to consider some fundamental aspects of experimentation, namely to evaluate the preferred measurements based on their feasibility and validity. In an ideal case, MOP/MOEs would be easy to measure (quantify), discriminate between control and experimental conditions, and contain a high level of external validity where evaluations would closely represent real life tasks.

As it pertains to feasibility, this report aims to evaluate the proposed metrics to acquire a general picture of whether (and where) performance measures can be implemented without significant difficulty. In this sense, difficulty refers to the level of effort needed to develop a study with respect to the time, resources, and level of complexity needed to administer and analyze research within the vVic environment. For example, a static target detection experiment measuring hit/miss percentages would be easier to execute than a study which aims to evaluate how well

operators are able to use Target Motion Analysis (TMA) to successfully define a “go deep” or “look” interval. In the latter case this would typically involve greater effort to execute, given the need to develop target attributes, (e.g., speed, sensors, size), scenario characteristics (e.g., target quantity and behaviour) and definition of ‘success’ (expert evaluation/confirmation).

Secondly, feasibility of research was considered in terms of the quality of information provided by the metric. As such, the intent was not to prioritize research solely based on ease of execution, but to balance the difficulty with how well MOP/MOEs can be applied in an externally valid context. Several suitable areas of evaluation were identified during the SME interviews and discussed later in Section 4.4. The current intent of this section is to acknowledge the mutual importance of feasibility and validity towards shaping future experimental plans, and to highlight a pragmatic approach to conceptualizing the various MOP/MOEs.

The following list provides a general view of measurements organized by their typical characteristics and ordered in terms of their anticipated difficulty (feasibility). These measures as described should not be regarded as rules without exception but rather as generic outlines intended for approximating a scale for the level of complexity that might be expected when planning, testing, or interpreting future research within the vVic:

- CATEGORY 1 – Objective, easily measureable without significant scenario development (e.g., isolated contact detect and classify scenario without environmental stimuli);
- CATEGORY 2 – Subjective metrics that can be standardized to evaluate performance and applied using an objective approach to scoring (e.g., behaviour anchors rated by an expert; a prescribed template for ‘good’ performance would likely be defined and consistently applied across all participants);
- CATEGORY 3 – Objective measures that require event driven context to necessitate an action due to consideration of two or more factors involving ‘trade-off’ decisions. May include several scenario phases and experimental markers to measure participant responses;
- CATEGORY 4 – Subjective metrics which are not anchored to behaviours tailored to the task and may apply more general HF concepts such as workload and SA (e.g., self-reports, probes, questionnaires, expert evaluation). These may also necessitate scenarios and require experimenters to mitigate confounds due to inter-rater reliability, and ensure that participants understand the nature of subjective reporting being requested (e.g., timeframe considered during post-hoc questionnaires, unnecessarily exuding over-confidence).
- CATEGORY 5 – Objective psycho-physiological metrics requiring special measurement apparatus that may be costly, intrusive, and time consuming. Increased effort to obtain approval from Ethics committees may also be a concern.

Measures of operator performance provide the necessary means to justify the user-centred design efforts. Performance is rarely easy to quantify at a gross level and often necessitates a conglomeration of both objective and subjective metrics to enable experimenters to profile task performance as it relates to effectiveness (success). The expected outcome of the interview sessions was the provision of a comprehensive set of MOP/MOEs which can be selectively tailored to align with future areas of experimentation. This is facilitated by the grouping of

metrics through a contextual understanding of the functional work areas identified during analysis. The next section provides a summary of the MOP/MOEs identified for future research.

4.3.3 Collection and Reduction of Data

The collection of measures identified in this section draw from a number of sources including the previous reports (Taylor et al, 2009; Bruyn Martin et al, 2009; Bruyn Martin et al, 2010; Bruyn Martin & Taylor, 2010; Rehak et al, 2011a; Rehak et al, 2011b), previous CAE contract work relating to submarines (Lamoureux et al, 2011), and the knowledge base provided through consultation with SMEs during the interview sessions. The approach to developing this list of MOP/MOEs was not to produce exhaustive content, but to gather the most observable forms of performance and effectiveness within relevant (representative) task groups exercised on the VICTORIA Class Submarine.

The full list of measures presented in this section are contained within an Excel spreadsheet and itemized across categorical groupings pertaining to the following considerations:

- Task Context – Groupings based on the work areas presented during the SME interviews (see 3.3) relating to the context under which tasks are conducted (e.g., ship handling, contact management, etc.);
- Difficulty Evaluation – Describes an estimated level of effort required to attain a MOP/MOE, typically due to the time, complexity, or resources required to conduct experimentation on the measure under consideration (see Section 4.3.2);
- Type of Metric – Refers to the nature of data measure, either as an objective (quantifiable) or subjective (e.g., opinions, rating scales, questionnaires) metric;
- Scope of Metric – Categorizes the listed measures as lower level metrics of task performance typically conducted by individuals (e.g., detect a signal, track a contact, steer true to course), or as higher level measures of system effectiveness (e.g., achieve mission, manage contacts, maintain covertness).

Table 7 provides an excerpt of the MOP/MOE list compiled during the course of this project (delivered as a separate Excel file). With the itemized organization, users of this spreadsheet can use Excel to sort on functions that may be of interest to a specific context. The example below shows a sort that has been applied to isolate the Contact Management measures, and ordered by increasing level of expected difficulty.

Table 7: Example MOP/MOE List Sorted by Contact Management and Difficulty Evaluation.

Metric	Difficulty / Evaluation	Type Metric	Task Context	Scope
Threshold: Angle on the Bow assessed within 5 - 10 deg accuracy	1	OBJ	contact mgmt	MOP
Threshold: Range estimation within 10% accuracy	1	OBJ	contact mgmt	MOP
Speed estimation based on range bearing readings (at least 2)	1	OBJ	contact mgmt	MOP
Was an accurate tactical picture maintained throughout the scenario (composite percentage of contacts detected/tracked and ID level)	1	OBJ	contact mgmt	MOE
Number of ‘lost’ contact incidences	1	OBJ	contact mgmt	MOE
Duration of time contacts were positively tracked vs duration of time contacts were out of contact	1	OBJ	contact mgmt	MOE
Number of contacts detected vs number in scenario	1	OBJ	contact mgmt	MOP
Number of ‘Unknowns’	1	OBJ	contact mgmt	MOP
Knowledge of current and future positions of all contacts in scenario	1	OBJ	contact mgmt	MOP
Number of COIs found and tracked (percentage)	1	OBJ	contact mgmt	MOP
Contact declarations to narrow threat class	2	SUBJ	contact mgmt	MOP
Appropriate selection of video size / quality	2	SUBJ	contact mgmt	
Contact priorities (divide into first, second, third/not a priority) (subject response to be evaluated by expert)	2	SUBJ	contact mgmt	MOP
Communication of contact priorities to crew (expert evaluation)	2	SUBJ	contact mgmt	MOP
Multi-tasking: maximize information gathering within a time window	3	OBJ	contact mgmt	MOE
Number of contact re-classification, false alarms, or repeated contacts	3	OBJ	contact mgmt	MOP
Was the mission successfully completed – yes or no?	3	OBJ	contact mgmt	MOE
Amplification level achieved for each/all contacts	3	OBJ	contact mgmt	MOP
Accuracy of TMA when compared to ground truth (course, bearing, range, speed)	3	OBJ	contact mgmt	MOP
Speed and accuracy of go deep/look interval calculations	3	OBJ	contact mgmt	MOP
Number of contacts that an officer can track for the purposes of look interval	3	OBJ	contact mgmt	MOP
Accuracy of predictions of collision threats	3	OBJ	contact mgmt	MOP
Completeness of tactical picture (expert evaluation)	4	SUBJ	contact mgmt	MOE
Confidence in understanding of tactical picture (self-report)	4	SUBJ	contact mgmt	MOE
Problem solving tradeoff between tracking contacts and dropping contacts (expert evaluation)	4	SUBJ	contact mgmt	MOE
Clarity of contact plot (i.e., no unnecessary stale, lost, time late contacts) (expert evaluation)	4	SUBJ	contact mgmt	MOE
Workload questionnaire/NASA TLX (self-report)	4	SUBJ	contact mgmt	MOE
Level of overemphasis on any element of the a priori information, intel or ORBAT (expert evaluation)	4	SUBJ	contact mgmt	MOP
Effectiveness of sensor employment (e.g., periscope vs towed array vs onboard sonar (active/passive) vs comms vs intel vs radar vs EW) (expert evaluation)	4	SUBJ	contact mgmt	MOP

4.3.4 Selecting Measures for Experimentation

The desire for simplicity in the proposed list does not preclude the need for a carefully selected set of MOP/MOEs to yield definitive, meaningful, and insightful results. In particular, the difficulty/evaluation approach to grouping measures does not suggest that future research should focus on 'CATEGORY 1' metrics based solely on economy of effort. Further, the relative evaluations of difficulty do not necessarily reflect a linear scale, and may not always be consistent across all use cases (e.g., CATEGORY 4 measures may not always be harder to achieve than CATEGORY 3 measures).

The proposed groupings and difficulty ratings are intended to help experimenters compile a large list of candidate MOP/MOEs and reduce them in such fashion that priority research areas can be aligned to fulfill project objectives. As such, it is recommended that a rationalized selection of measures be developed by drawing from the existing list of MOP/MOEs and identifying candidate experimental cases under which they can be most appropriately applied. Researchers should be mindful of the cost-benefit associations between experimentally controlled testing (e.g., isolating effects of independent variables) and the improved external validity of complex, but less controlled experimental scenarios. The interview sessions with the SMEs provided valuable insight and validation on how to approach candidate use case scenarios for experimentation. These considerations are discussed next.

4.3.5 Cross Functional Considerations

In review, the original approach of having SME participants individually rate a list MOP/MOEs for suitability was re-oriented to focus on a more collective discussion of the measurement areas of interests (listed in Section 3.2). This approach was selected due to the time available (more expedient) as well as to better leverage the tactical expertise provided by SMEs familiar with the CO/OOW positions. Further, the background of SME participants was regarded as more fitting to discuss performance and effectiveness measures from a broader (e.g., tactical considerations) sense rather than to rate them based on experimental feasibility.

During the interviews, it was evident that many of the functional work areas presented to the SMEs were not evaluated in isolation, but as joint factors that contributed to an overall picture of 'success'. In general, the SME projections of higher level goals was consistent with those reviewed from the CWA reports (Taylor et al, 2009; Bruyn Martin et al, 2009; Bruyn Martin et al, 2010; Bruyn Martin & Taylor, 2010; Rehak et al, 2011a; Rehak et al, 2011b), where Safety, Covertness, and Mission was cited[2] as the overarching goals that were likely to influence operator behaviour in terms of decision making, task selection and prioritization. Cross functional (combined) considerations were hence the required context needed for the SMEs to answer questions regarding the most suitable metrics for evaluating performance and effectiveness. Although significant objections to presentation materials were not raised by the participants, the objectives of this project necessitated a clear connection between measurements and the scenario context (e.g., events and environmental information) before researchers could specify individual MOP/MOEs. Through discussions with the SMEs, a number of examples were proposed to illustrate how MOP/MOEs would be jointly considered and the circumstances under which they

[2] Officially termed, in order of priority: 1) Safety of the submarine and her personnel, 2) Remaining undetected, and 3) Achieving the aim.

would need to be evaluated (e.g., trade-off or prioritized). These discussions are reviewed in Section 4.4.

The presentation of information collected during this project involves the identification of candidate groupings of work as suggested by the SME participants. This revised[3] framework demonstrates an operationally plausible grouping of work tasks that can be developed for experimentation. Within these constructs, researchers can extend work by applying variable degrees of complexity based on the anticipated level of effort (e.g., project scope and budget). The following sections present this approach as packages of research, comprised of appropriated measures of performance and effectiveness based on the functional work areas relevant to the situation (mission).

4.4 Experimental Vignettes

The combination of contextual task groupings and the list of MOP/MOEs form a number of experimental vignettes that can be used to guide future areas of experimentation. These vignettes are not intended to complete the details needed to enact scenarios for experimentation, but to highlight key aspects of performance and effectiveness within operationally relevant areas of conduct for the CO/OOW and crew onboard the VICTORIA Class Submarine. Each of the vignettes can be applied as either isolated (individual) tests, or as phases of an extended mission scenario involving several factors, depending on the scope and objectives for future experimentation. To assist with abstracting level of effort, the suggested measures for each vignette have been categorized using the following structure:

- Vignette Overview – Summarizes the nature of discussion from SME interviews, performance factors, or scenario events that would be used to characterize each of the proposed vignettes. The general purpose of the vignette is described in terms of how performance and effectiveness measures would be elicited within the setting of each vignette;
- Basic Measures – A selection of the most readily achievable MOP/MOEs suitable for the vignette under consideration. These measures are drawn from the list presented in Annex A and shall correspond with the context of the vignette presented. The basic measures will typically match up with a Difficulty / Evaluation rating of 1 or 2, but not without exceptions. Basic measures attempt to capture the relevant aspects of the operator work tasks within each setting and to delineate sample metrics of performance without significant scenario development. Several basic measures may be required to arrive at an overall definition of effectiveness (i.e., MOE). Basic measures may be either subjective or objective.
- Composite Measures – A selection or set of MOP/MOEs that necessitates the development of greater scenario complexity or environmental representation to stimulate multi-tasking, task trade-offs, or unknown information gathering tasks. While more effort is expected to achieve these measures, a significant amount of insight can be acquired through accurate representations of task demands involving multiple concurrent factors. Accurate representations of complex environments are of key importance towards

[3] Does not imply that the work areas presented during the SME interviews were grouped incorrectly, but highlights that they are often jointly considered and not in isolation.

attaining accurate measures of the subjective areas, particularly as they relate to situation awareness and workload. The basic measures will typically match up with a Difficulty / Evaluation rating of 3 or 4 but not without exceptions. These measures are similarly drawn from the list presented in Annex A and shall correspond with the context of each vignette. Composite measures may be either subjective or objective.

- IID Measures – A selection or set of MOP/MOEs that focus specifically on the design of the IID and its contribution to the tasks of the command team aboard the VICTORIA Class. The list elements have not been subjected to a consideration of their difficulty to evaluate. The measures listed under this heading are not necessarily the only MOP/MOEs that may shed light on the utility and usability of the IID; MOPs and MOEs intended for work functions may also support the consideration of IID utility and usability.

The intent of the experimental vignettes is to provide a sufficient sampling of work tasks and trigger appropriate behaviours and which might be common across the majority of mission profiles. Further validation of these vignettes with SMEs should be exercised when conducting future research. The following sections are examples of the cross consideration of functional work areas introduced previously, and describes the selection of metrics within contextual frameworks:

4.4.1 Navigation to Destination Vignette

4.4.1.1 Navigation to Destination Overview

The Navigation to Destination vignette was brought forth during SME discussions related to ship handling and the comparison between planned routes and adherence to course. It was suggested that in order to elicit performance during navigation, researchers should consider events that would require crews to deviate from course and consider factors outside of basic route maintenance. The implication was that fundamental ship handling duties such as maintaining true to course should not pose a significant challenge to trained operators, assuming fully functional submarine systems and a non-threatening environment (e.g., safety, environmental states, etc.). This suggests that in order to use ship handling as an indicator of performance, factors which influence the overall conduct of navigation should be represented during experimentation. In general, these factors are the tactical considerations during the execution of planned routes that influence the overarching goals for the ship (e.g., safety, covertness, or mission) and consequently may take precedence over current ship handling activity.

The considerations determined to have a high likelihood of impacting crew behaviour within navigation contexts include:

- Tactical Utilization of Environment – Ship manoeuvre and navigation activity (either deliberated or reactionary) along the existing route that crews might employ to achieve a tactical advantage. This may involve the allocation of intermediary waypoints along a route to enhance covertness (probability of detection), general maintenance (or avoidance) of depths, vicinity to land, and obscuring own ship signatures among other vessels, depending on the fidelity of the environmental simulation available;
- Avoidance of Blind Sensor Areas – Ship handling activity or patterns of manoeuvre that reduce the time spent without sensor coverage in particular bearings or areas relative to the ship;

- Surveillance of Mission Critical Areas – Ship positioning or patterns of movement that maximize the crews ability to maintain observation of mission critical areas (through the periscope or sensors). These tactics often involve the consideration of ship sensor capabilities and their limitations (depth and range) in order to select the most appropriate course of action. This pursuit can be considered as a counterpart to the previous itemized activity (avoiding blind spots);
- Recovery of Tactical Advantage – Ship manoeuvring in response to a counter-detection incident by another unit while navigating along a route. The general goal of maintaining covertness provides a tactical advantage that submarines consistently intend to exploit. Depending on threat level of the unit which has counter-detected, the submarine may need to deviate from course either to perform evasive manoeuvre or to regain the advantage of covertness; and
- Economy of Ship Resources – The effective utilization of ship resources such as air, battery power, and fuel which maximizes efficiency in terms of an appropriate balance between the time taken to reach a destination and the resources consumed to do so.

4.4.1.2 Navigation to Destination Basic Measures

The basic measures selected for the Navigation to Destination vignette are presented below. These metrics were drawn primarily from the Covertness, SA, Planning, and Ship Handling task contexts as presented during SME discussions (Section 3.2) and may incorporate further insights collected during the interviews. The categorical information pertaining to the Type (OBJ/SUBJ) and Level (MOP/MOE) as outlined in the MOP/MOE Excel list (0) are also presented:

- Time spent at periscope depth (OBJ, MOP) – Assess depth selection during navigation. This measure may be re-termed or grouped with other depths of relevance to the mission (e.g., snorkel depth);
- Battery status (perception level 1 SA) (OBJ, MOP) – An individual SA measure that can be used to identify current crew awareness of battery power. This may be used to assess subsequent crew behaviour or decision making based on the mission profile;
- Speed, depth, course of submarine (perception level 1 SA) (OBJ, MOP) – An individual SA measure that can be used to identify current crew awareness of ship status information related to navigation. This may be used to assess subsequent crew behaviour or decision making based on the mission profile;
- Completeness of information contained in navigation plan (e.g., speed, heading, leg duration, depth, snorting interval, look interval, communications, engineering, life support, domestics, sensor use, weapons use) (SUBJ, MOP) – A potential checklist of information (e.g., behaviour anchors) that can be used to assess performance during navigation planning.

4.4.1.3 Navigation to Destination Composite Measures

The composite measures selected for the Navigation to Destination vignette are presented below. These metrics were drawn from the Excel list and derived primarily from the Covertness, SA, Planning, and Ship Handling task contexts:

- Accuracy of synchronization with other assets (OBJ, MOE) – Evaluate overall success of navigation pending mission profiles requiring synchronization activity (e.g., rendezvous at waypoints, avoid crossing border boxes, etc.);
- Adherence to boxes (friendly boundaries, routing, etc.) (OBJ, MOP) – Supplementary to the previous MOE as a constituent of performance contributing to effective synchronization with assets; can be measured as instances when the ship crosses boundaries that it should not (determined through experimental scenario);
- Successful evasion (OBJ, MOE) – Used to assess the crews' ability to regain covertness in the event of counter-detection (as represented during scenarios). The definition of evasion may need further refinement through SME consultation (e.g., achieve depths, manoeuvre outside of detection range);
- Number of counter-detections (OBJ, MOE) – Supplementary to the previous MOE as a constituent of performance contributing to effective evasion; can be triggered through scheduled scenario events or instances when the ship enters within the detection range of a contact (determined through experimental scenario);
- Percentage of time observing or not observing (blind) mission critical areas (OBJ, MOP) – A measurement of effective sensor application and utilization during navigation. The metric would generally assess the operators' ability to position or manoeuvre the ship in order to effectively maximize and exploit sensor ranges and to reduce occurrences of blind coverage;
- Comparison of submarine historical track with the planned (deviation) (OBJ, MOE) – A visual representation or percentage of time spent off track to assess adherence to course. This metric is subject to interpretation regarding success as it can be influenced by several factors where deviation from course is not necessarily undesirable (e.g., safety, covertness);
- Duration of actual mission vs. planned duration (OBJ, MOP) – Supplementary to the previous metric regarding course deviation as an individual performance metric. This is also subject to interpretation as expediency may not always take precedence;
- Degree to which any deviations were accounted for in contingency planning (expert evaluator) (SUBJ, MOE) – Expert ratings of whether evaluation participants took the necessary precautions or considerations prior to navigation to allow crews to anticipate the need to deviate from course. Typically these are scenario driven and rated in response to events;
- Degree to which mission objectives were being satisfied, including safety and covertness (expert evaluation) (comprehension and project, level 2 & 3 SA) (SUBJ, MOE) – Expert ratings of whether evaluation participants took the necessary information into consideration to arrive at decisions or courses of action. These subjective ratings require well defined scenarios that can be tailored to allow ratings of 'correct' actions in response to events during the scenario (e.g., a standardized rating system to maintain good reliability);
- Successful execution of navigation plan (SUBJ, MOE) – A subjective rating likely to be defined by a desired list of objectives pertinent to the mission. These may be comprised of task level MOPs to build up a rating of success and potentially applied as a score (e.g., percentage of navigation objectives achieved).

4.4.1.4 IID Related Measures

There are a number of elements of the IID that could assist the navigation task. The difficulty is in measuring the SME's use of the elements. To this end, there are a number of standard MOPs and MOEs that should be applied to the IID.

- EMT Scan Patterns (OBJ, MOE) – the pattern by which the SME scans the IID to carry out tasks associated with navigation should be studied and compared with the scan patterns (including verbal or audible information acquisition tasks) of current navigation tasks. Any differences should be noted, especially if the SME begins to omit information sources when using the IID. This latter observation may either be evidence of good design (i.e., that the information is presented or combined implicitly with other IID elements) or over-reliance on particular information sources. With familiarity scan patterns should become much more direct and less 'dense' with respect to the number of pauses.
- EMT Dwells (OBJ, MOP) – the time (percentage of time spent looking at IID, percentage of time overall) spent looking at specific IID frames should be studied and compared with the amount of time spent looking or attending to analogous information sources in the current control room. In particular, good design should result in fewer dwells of shorter duration, especially as familiarity with the system grows.
- Mouse Movements (OBJ, MOP) – similar to EMT Dwells above, this would measure the time (percentage of time spent interacting with IID, percentage of time overall) spent interacting with specific IID frames, and compared with the amount of time spent interacting with analogous information sources in the current control room. Good design should result in few interactions required for the SME to reach their goal, especially as familiarity with the system grows.
- IID Configuration (OBJ, MOE) – for navigation purposes the IID would typically show the Overall Tactical Picture, supplemented by complementary views in the dynamic frames. The views used by the SME, as well as what they switch to in which frame, should be tracked and studied. If the SME has to switch views frequently, this may indicate a poor design of the IID.

Additionally, there are a number of specific MOPs and MOEs that can be applied to the IID.

- Accuracy of Navigation (OBJ, MOP and MOE) – the timings and course of the submarine should be compared between IID and non-IID conditions. A well-designed and useful IID will result in better adherence to the planned course and punctuality with respect to planned timings.
- Time to Evasion (OBJ, MOP and MOE) – the time taken to successfully regain covertness after a counter-detection should be compared between IID and non-IID conditions. A well-designed and useful IID will result in a better appreciation of the evasion options, quicker decisions, and an earlier evasion.
- Effectiveness of Depth Selection (Expert Evaluator) (SUBJ, MOP) – the IID includes a Sound Velocity Plot. This should assist the SME to select the best depth at which to drive the submarine while maximizing the covertness of the submarine. The expert evaluator could rate depth selection on a scale of 1 to 5 for the whole simulation, and then results can be compared between IID and non-IID conditions.

Additionally, most of the basic and composite measures described above can be used in an IID/non-IID comparison to infer the impact of the IID.

4.4.2 Contact Management Vignette

4.4.2.1 Contact Management Overview

The Contact Management vignette provides several potential interactions that can be considered within a variety of tactical situations. The process of Contact Management refers to the initial detection, identification and tracking, through to the prosecution[4] of contacts (if required). Tasks may include the correlating of detection information with intelligence data, as well as the effective use of sensors to gather more information on the contact for correlation with a priori data (e.g., Order of Battle [ORBAT]). Decision making regarding collision threats, go deep levels, look intervals, contacts of interest (COI), and closest points of approach (CPA) are of key interest to maximize effectiveness. Effective problem solving (e.g., TMA) for tracking and surveillance are the common areas of work used to arrive at solutions during Contact Management.

The following list of scenario characteristics was identified for representation in future research following discussion with the SMEs:

- Threat Reduction Process – The narrowing classification of contacts based on a fundamental assumption that all contacts arise as high threats. Sufficient information regarding contact details should be presented during experimentation in order to exercise the threat reduction strategy typically employed by operators. This strategy may vary between crews, but generally involves a form of high level grouping of class (e.g., military / civilian), determination of contacts intent (e.g., fishing, searching, unknown), and the expected number and type of units operating within an area based on information provided by higher intel;
- Contact Density – Efficiency in contact classification is important to crews but not always achievable within contact rich environments. Workload burden can be a crucial driver of performance and should be assessed during experimentation. The SMEs indicated that a situation involving 20 – 25 contacts requiring TMA would be considered a busy situation.
- Tactical Picture Maintenance – The management of contact information in order to develop an accurate, up to date tactical picture without significant clutter or incorrect information leading to mistakes. The potential forms of error can be measured in terms of contact duplication, instances of re-classification, and missed priority contacts;
- Contacts of Interest – High priority targets that often result in critical task execution or special attention from the crew. These contacts would likely be framed within a situational context such as evading contacts (e.g., an enemy ship hiding among fishing vessels), surveillance and identification along drug routes, or avoidance of high threat contacts (submarines, warships) and counter-detection;

[4] Target prosecution and weapons employment are outside the scope of this project and not discussed in further detail from this point forward.

- Safety, Covertness, Mission Tradeoffs – The management and balance between risk taking and mission objectives can be exercised in a scenario where two or more high priority goals are in conflict and crew effectiveness can be compromised. The aim of evaluation would be to measure crew performance under stress, such as exposure to multiple high priority, time critical decisions involving numerous factors. Within the contact management vignette, this could result in crews having to focus on priority contacts at the expense of losing other contacts which can be increasingly challenging in busy environments.
- Depth Selection or Maintenance – A critical area of consideration during contact management are tasks associated with visual surveillance, tracking, approach and prosecution. Depth management is a definite case where safety, covertness, and mission tradeoffs would occur and can be tailored into experimental scenarios for assessing performance (e.g., situations requiring the sub to go deep, or return to periscope depth).
- Target Motion Analysis – Interface level solutions to facilitate operator tasks as they pertain to workload and situation awareness. During interviews, some SMEs proposed the utility of sonar overlays when conducting TMA to increase effectiveness when developing target solutions. Specific evaluations pertaining to the TMA interface and functionality may be possible through the IID development project at DRDC Atlantic.

4.4.2.2 Contact Management Basic Measures

The basic measures selected for the Contact Management vignette are presented below. These metrics were drawn from the Excel list and derived primarily from the Contact Management and SA task contexts:

- Discrete estimation of contact information regarding:
  - Range (OBJ, MOP) – An acceptable threshold for range estimations were identified to be within 10% accuracy based on SME feedback. Representations of contacts within the simulation would be required to establish the true range to compare against operator estimations.
  - Angle on the Bow (AOB) (OBJ, MOP) – Acceptable limits for AOB estimations were identified to be within 5 – 10 degrees of accuracy based on SME feedback. Representations of contacts within the simulation would be required to enable AOB estimation scenarios.
  - Manual target motion analysis such as speed, bearing, and heading (OBJ, MOP) – Speed estimation requires a minimum of two range-bearing points (representations) at discrete points in time to establish. This metric allows experimenters to assess the operators' ability to conduct manual estimations without assisted resolution of target motion (e.g., TMA).
- Discrete determination of contact information regarding:
  - Number of contacts detected (and missed) (OBJ, MOP) – This measure can be expressed in terms of COIs through scenario tailoring; typically there is surrounding noise that operators would need to disregard based on target

prioritization rules. The absence of noise while searching or surveying potential contacts (e.g., via radar, acoustic, periscope sensors) would yield less meaningful information because of low difficulty and inaccurate portray of actual task demands.

- Number of contacts classified correctly (and incorrectly), including unknown contacts (OBJ, MOP) – Contact classifications are often verbalized through contact 'declarations' as a part of crew protocol. Verbal declarations such as these may require measurement in stages (several declarations for each contact) and likely necessitates the need to tailor verbal scripts during experimentation for measurement consistency.
- Number of duplicate contact detections (OBJ, MOP) – Instances of repeated contact detections can be used as an indicator of workload burden leading to operator error. The measurement protocol on how to record these instances are important for purposes of measurement consistency and reliability.
- Communication of contact priorities to crew (expert evaluation) (SUBJ, MOP) – A subjective measure of SA that can be used to investigate successful prioritization (or confusion) among crew members. The fundamental factors which influence the level of threat classification and resulting priority should be validated with SMEs before developing a scoring approach.

4.4.2.3 Contact Management Composite Measures

The composite measures selected for the Contact Management vignette are presented below. These metrics were drawn from the Excel list and derived primarily from the Contact Management and SA task contexts:

- Accuracy of TMA when compared to ground truth (heading, bearing, range, speed) (OBJ, MOP) – TMA activities have significant relevance and impact to performance during contact management, however the task level interactions with TMA displays need to be represented in order to enable measurement within this area;
- Amplification level achieved for each/all contacts (OBJ, MOP) – Measures the degree to which preliminary contact information can be enhanced and consolidated, particularly for COIs in order to determine ongoing priority and courses of action. Amplification task representations will vary depending on the fidelity of simulation available;
- Speed and accuracy of go deep/look interval calculations (OBJ, MOP) – Measures the ability of crews to use contact information to execute appropriate courses of action. Within this context, the target capability would determine appropriate look intervals or range at which the submarine should dive to a certain depth. Predictions are typically calculated based on time factors related to speed of contacts and the safe range at which to avoid being hit by a threatening vessel. The CPA can be applied for this purpose;
- Frequency and interval of all round looks (OBJ, MOP) – A measure which can be used as an indicator to determine when crews are challenged to find a surface contact. All round looks are generally used when the position of a surface contact is uncertain, and visual confirmation is required;

- Bearing rate tolerance (within 'x' meters per second) (OBJ, MOP) – Situational measure of performance that can be used to evaluate operators' ability to balance safety considerations (arising from contacts) against other pressing objectives (e.g., snort to mitigate poor air quality);

- Completeness and Clarity of tactical picture (expert evaluation) (SUBJ, MOE) - A subjective scoring of tactical picture compilation to determine if sufficient and appropriate information was gathered for the COIs; the presence of unnecessary information (stale, duplicate, time late contacts) that may obscure the tactical picture could also be used to assess performance in these areas;

- Problem solving trade-offs between tracking contacts and dropping contacts (expert evaluation) (SUBJ, MOE) – Evaluates the decision making competency of crews pertaining to prioritization or recognition of key information aspects particularly during situations involving high workload (e.g., dense contact environment);

- Effectiveness of sensor employment or selection (e.g., periscopes, sonar, radar, EW, etc.) (expert evaluation) (SUBJ, MOP) – A situational measure of performance to determine whether operators were able to appropriately use (and exploit) sensor capabilities based on the contacts existing within a scenario. This metric is subject to further SME validation to determine if a definitive method of sensor employment can be standardized for specific target profiles or situations (e.g., when to avoid periscope use, when to avoid active transmissions);

- Shared SA regarding priority contacts (subject response to be evaluated by expert) (comprehension level 2 SA) (SUBJ, MOP) – A subjective critique of the factors that were considered when determining the relative priorities of contacts within a scenario. This measure offers insight pertaining to how SMEs arrive at decisions and demonstrates how workload may lead to confusion and failure to notice critical information;

- Shared SA regarding Commander's intent (subject response to be evaluated by expert) (comprehension and projection level 2 & 3 SA) (SUBJ, MOP) – The assessment of crew coordination and understanding of individual task level responsibilities as they relate to the overarching goals of the ship (safety, covertness, mission). The impact that individual tasks have on overall objectives may not be readily apparent, but necessary in order for the crew to have shared awareness of the Commanders intent. While this metric is abstract, it offers potential to identify critical information requirements along the chain of command where design interventions or procedural modifications can improve performance;

- Frequency of updates pertaining to the submarine environment (comprehension and projection level 2 & 3 SA) (SUBJ, MOP) – A subjective assessment regarding the influence of Commander updates (sit reps) on the overall effectiveness towards contact management. Situation Reports (sit reps; detailing current activities, progress, priorities, situation, environment, etc.) can vary in duration and level of complexity but generally includes knowledge of mission progress, future plans, crew status, and submarine health. This measure can also be used towards investigating the degree to which unsolicited information can impact crew attention (or distraction) towards critical information. During the interviews the SMEs emphasized that shared SA is a process of optimal information distribution rather than communication for sake of volume or frequency.

4.4.2.4 IID Related Measures

There are a number of elements of the IID that could assist the contact management task. The difficulty is measuring the SME's use of the elements. To this end, there are a number of standard MOPs and MOEs that should be applied to the IID.

- EMT Scan Patterns (OBJ, MOE) – the pattern by which the SME scans the IID to carry out tasks associated with contact management should be studied and compared with the scan patterns (including verbal or audible information acquisition tasks) of current navigation tasks. Any differences should be noted, especially if the SME begins to omit information sources when using the IID. This latter observation may either be evidence of good design (i.e., that the information is presented or combined implicitly with other IID elements) or over-reliance on particular information sources. With familiarity scan patterns should become much more direct and less 'dense' with respect to the number of pauses.
- EMT Dwells (OBJ, MOP) – the time (percentage of time spent looking at IID, percentage of time overall) spent looking at specific IID frames should be studied and compared with the amount of time spent looking or attending to analogous information sources in the current control room. In particular, good design should result in fewer dwells of shorter duration, especially as familiarity with the system grows.
- Mouse Movements (OBJ, MOP) – similar to EMT Dwells above, this would measure the time (percentage of time spent interacting with IID, percentage of time overall) spent interacting with specific IID frames, and compared with the amount of time spent interacting with analogous information sources in the current control room. Good design should result in few interactions required for the SME to reach their goal, especially as familiarity with the system grows.
- IID Configuration (OBJ, MOE) – for contact management purposes the IID would typically show the Overall Tactical Picture, supplemented by the contact management map or table to the left side and current contact list or totes in the frame below. The views used by the SME, as well as what they switch to in which frame, should be tracked and studied. If the SME has to switch views frequently, this may indicate a poor design of the IID.

Additionally, there are a number of specific MOPs and MOEs that can be applied to the IID.

- Amplification level achieved for each/all contacts (OBJ, MOP) – amplification levels for each contact and overall should be compared between IID and non-IID conditions. Because of improved decision making and information presentation afforded by the IID, leading to better SA, the expectation should be that the IID results in a higher degree of average identification amplification.
- Speed and accuracy of go deep/look interval calculations (OBJ, MOP) – the speed with which the SME can provide go deep/look interval information, as well as the accuracy of that information, should be obtained (by experimenter's query) and compared between IID and non-IID conditions. Also, the SME's ability to track these intervals accurately should also be tracked. Although the IID might not result in quicker recall of this information, the accuracy should be greater, as well as the SMEs ability to accurately follow the calculated timings.

- Frequency and interval of all round looks (OBJ, MOP) – similar to the go deep/look interval calculations above, the number and interval between all round looks should be tracked and compared between IID and non-IID conditions. Use of the IID should result in more efficient application of all round looks, possibly culminating in a reduced requirement if the IID presents accurate information about contacts.
- Bearing rate tolerance (within 'x' degrees per second) (OBJ, MOP) – bearing rate tolerance between IID and non-IID conditions should be compared. The IID should result in greater accuracy and thus safety and covertness.

Additionally, most of the basic and composite measures described above can be used in an IID/non-IID comparison to infer the impact of the IID.

4.4.3 Window of Opportunity Vignette

4.4.3.1 Window of Opportunity Overview

The Window of Opportunity vignette provides a framework of measurement that corresponds with general surface or near-surface activity (e.g., periscope depth) that encapsulates a number of tasks pertaining to maintenance of boat health and the enabling of communications. During the interviews, the SMEs commented that crew proficiency may be difficult to pinpoint under routine circumstances and suggested experimentation during situations involving greater time pressure. The discussion specifically referenced surface scenarios, where multiple tasks were of the essence and timing was crucial to reduce submarine exposure (safety).

Crew performance degradations would accordingly be the expected outcome during situations of high workload and time urgency. As such, this vignette allows experimenters to set a benchmark for performance under varying conditions of workload and time pressure (e.g., baseline workload, high workload, etc.). It should be noted that the vignette described in this section deliberately omits surface activity related to high priority (threat) contact management (see Section 4.4.2) or emergency / incident response taskings (see Section 4.4.4). The intent of the Window of Opportunity is to capture the less urgent surface tasks that are ‘opportunistic’ and potentially shed at the expense of higher priority duties.

Scenarios can be subsequently tailored to incrementally challenge operators through a series of concurrent events which require operator attention and prioritization (e.g., contact management in Section 4.4.2; incident response in Section 4.4.4). This measurement vignette proposes that crew proficiency can be measured through task efficiency during multi-tasking situations; under the same time window of opportunity, researchers would expect poorer task performance with increasing workload challenges. The following series of work tasks can be used for developing the Window of Opportunity vignette:

- Maintaining the air quality – The general maintenance of the four primary gases in the air that crews breathe: nitrogen, oxygen, argon, and carbon dioxide. Air must be kept breathable within the submarine which requires surfacing in order for oxygen to be replenished and the removal carbon dioxide and exhaled moisture. Within the vVic, these considerations would be appropriate tasks to incorporate during a snorting scenario, where air replenishment would be among several tasks that need to be attended to. Interface representations would be required to simulate these interactions;

- Maintaining the fresh water supply - Most submarines have a distillation apparatus that can take in seawater and produce fresh water. This water is used mainly for cooling electronic equipment (such as computers and navigation equipment) and for supporting the crew (for example, drinking, cooking and personal hygiene). While this area of work may not exist within the vVic, a simplistic representation of this information on the operator interfaces will be useful to better reproduce considerations related to maintenance of internal environment and life support (i.e., safety goals for the submarine);
- Maintaining the temperature - The temperature of the ocean surrounding the submarine is typically lower than that of its internal environment, so submarines must be electrically heated to maintain a comfortable temperature for the crew. Temperature maintenance (and associated battery depletion) provides an additional reason for crews to engage in surfacing activity to obtain the oxygen needed to power the diesel-electric engine of the VICTORIA Class Submarine;
- Maintaining the power supply – There are several sources of equipment on board the VICTORIA Class Submarine requiring electric power. To supply this power, the VICTORIA Class Submarine is equipped with diesel-electric engines that burn fuel to charge the batteries. This process requires the submarine to at minimum cruise just below the surface (snorkel) in order to obtain the oxygen needed to enable combustion. Simulation of power draw and depletion can likely be represented on the vVic without significant development effort. Power supply considerations are critical to the health and overall effectiveness of the submarine and should be considered during performance evaluations;
- Communications with other assets or higher command – Communications are typically limited during deep submersion and may require the submarine to return to periscope depth to exchange information. Scenarios involving communication between the SUBOPAUTH and vVic are valid instances of activity near the surface but may involve small time windows to execute the set of tasks that require attention. Extended operations near the surface are generally considered less safe and therefore undesirable due to risk of compromised submarine safety. Task performance during time pressure can be a valuable vignette under which to measures performance as these may yield errors and reveal areas where design solutions can alleviate task challenges (e.g., improve efficiency, optimize information displayed, task bottlenecks, etc.);
- Surface contact management – Generally refers to the tactical tasks involved with managing surface vessels. These may include look intervals or all round looks to gather contact information and the environment in which ships are operating. This vignette does not dwell on the contact management aspects to reduce overlap (discussed in Section 4.4.2).

4.4.3.2 Window of Opportunity Basic Measures

The basic measures selected for the Window of Opportunity vignette are presented below. These metrics were from the Excel list and derived primarily from the Communications, Covertness, Safety and SA task contexts:

- Duration of communication (OBJ, MOP) – The time elapsed during an instance of communication (e.g., radio). Within the Window of Opportunity context, the duration required to complete communications would likely involve a pre-defined script and set of necessary information to be gathered within a scenario (e.g., tailored sit reps). This ensures reliability of the measurement;
- Response times of information requests (OBJ, MOP) – Time elapsed for short burst information requests or updates. This metric can be used as a sub-set of information within a collection of communication metrics to assess efficiency;
- Frequency of communication (OBJ, MOP) – The number of occurrences for communication can be used as an indicator of crew workload, SA, and effective performance. This will depend on the structure of the communication scenario and involve behaviours of interest such as clarification requests, or instances of corrective action by the CO / Watch Leader. This metric can be used as a sub-set of information within a collection of communication metrics to assess quality and accuracy;
- Time spent at periscope depth (OBJ, MOP) – Assess depth selection during a window of opportunity. This measure may be re-termed or grouped with other depths of relevance to the mission (e.g., snorkel depth);
- Look intervals (OBJ, MOP) – Assess operator allocation of time for surface search. This measure may be viewed in conjunction with the time allocated towards conducting other tasks to determine crew performance in terms of prioritization and efficiency;
- Declare course of action (expert evaluation) (SUBJ, MOP) – For simulations capable of representing boat health on crew interfaces, a subjective assessment can be used to determine how effectively they were able to respond in terms of priority and efficiency;
- Appropriation of time spent on tasks (expert evaluation) (SUBJ, MOP) – An expert rating considering the various allocation of time towards work tasks and determining if they were in line with the highest priority tasks. A predetermined mission scenario would dictate which tasks should be of greater priority and consequent weighting of time.

4.4.3.3 Window of Opportunity Composite Measures

The composite measures selected for the Window of Opportunity vignette are presented below. These metrics were drawn from the Excel list and derived primarily from the Communications, Planning, Covertness, Safety and SA task contexts:

- Number of milestone tasks completed (OBJ, MOE) – The quantity of work functions that were completed within a specified surfacing event (e.g., snort, collected contact information, communication transmissions). The work tasks should be derived from an experimenter defined list of important objectives as guided by SME assisted scenario development;
- Appropriateness of Communication (expert evaluation) (SUBJ, MOE) – Subjective ratings pertaining to the timeliness and necessity of messages during communication scenarios. These may include debriefings or expert opinions addressing the specific factors and events which led subjects to instances of compromised communication (whether deliberate or not);

- Command team was in possession of necessary information and shared all necessary information (expert evaluation) (SUBJ, MOE) – A grouping of subjective expert ratings to determine the succinctness and sufficiency of information sharing between the crew members;
- Elegance of planning for a predetermined window of opportunity (expert evaluator) (SUBJ, MOE) – The successful execution and attainment of objectives for pre-planned activities during an expected surface (or near surface) event, allowing for smooth transitions from one planned milestone to the next. This generally refers to the extent to which a plan was followed and the degree to which deviations were accounted for in contingency planning.

4.4.3.4 IID Related Measures

There are a number of elements of the IID that could assist the window of opportunity task. The difficulty is measuring the SME's use of the elements. To this end, there are a number of standard MOPs and MOEs that should be applied to the IID.

- EMT Scan Patterns (OBJ, MOE) – the pattern by which the SME scans the IID to carry out tasks associated with window of opportunity planning and execution should be studied and compared with the scan patterns (including verbal or audible information acquisition tasks) of current window of opportunity tasks. Any differences should be noted, especially if the SME begins to omit information sources when using the IID. This latter observation may either be evidence of good design (i.e., that the information is presented or combined implicitly with other IID elements) or over-reliance on particular information sources. With familiarity scan patterns should become much more direct and less 'dense' with respect to the number of pauses.
- EMT Dwells (OBJ, MOP) – the time (percentage of time spent looking at IID, percentage of time overall) spent looking at specific IID frames should be studied and compared with the amount of time spent looking or attending to analogous information sources in the current control room. In particular, good design should result in fewer dwells of shorter duration, especially as familiarity with the system grows.
- Mouse Movements (OBJ, MOP) – similar to EMT Dwells above, this would measure the time (percentage of time spent interacting with IID, percentage of time overall) spent interacting with specific IID frames, and compared with the amount of time spent interacting with analogous information sources in the current control room. Good design should result in few interactions required for the SME to reach their goal, especially as familiarity with the system grows.
- IID Configuration (OBJ, MOE) – for incident response purposes the IID would typically show the Overall Tactical Picture, supplemented by the schedule of events on the left and the primary ownship status above. The views used by the SME, as well as what they switch to in which frame, should be tracked and studied. If the SME has to switch views frequently, this may indicate a poor design of the IID.

Additionally, there are a number of specific MOPs and MOEs that can be applied to the IID.

- Alerts (OBJ, MOE) – the IID presents a number of configurable alerts, at least some of which should deal with air quality, battery power, fuel level, communications plans, etc. The objective would be to create a plan and execute it that results in the alerts never occurring. The number of alerts that occur should be recorded and compared with analogous data in the current control room. If the IID is well-designed, there should be fewer alerts or similar erosions of safety in the IID condition compared to the non-IID condition.
- Exposure of Periscope and EW mast (OBJ, MOP) – because the IID has a periview it has been assumed the periscope's exposure will be controlled automatically rather than manually. This should lead to reduced exposure of the mast to obtain an equal level of clarity for the scene. A comparison of IID and non-IID conditions should be made to investigate this.

Additionally, most of the basic and composite measures described above can be used in an IID/non-IID comparison to infer the impact of the IID.

4.4.4 Incident Response Vignette

4.4.4.1 Incident Response Overview

The Incident Response vignette was developed primarily to capture task groupings that are not deliberate but reactive in nature, often necessitating an immediate response or re-prioritization due to high priority or emergency events. These events are generally associated with a sudden compromise to boat safety either within the internal environment (e.g., gas leak, fire, smoke) or as an external threat (e.g., high threat contact, counter-detection, torpedo attack).

The Incident Response vignette provides researchers the opportunity to impose decision making tradeoffs under varying degrees of time pressure by overlaying priority events onto existing task conduct. The general prospective for this vignette would be similar to the Window of Opportunity (see Section 4.4.3), where scenario incidents would compete for the crews' attention based on urgency and importance to mission. The measurement aspects of interest involve their ability to prioritize tasks and determining where design interventions might improve task performance (e.g., expediency, reduced task shedding/errors, and multi-tasking).

As such, it is prudent to apply this vignette in a comparative experimental construct where difficulty is applied in incremental fashion (e.g., baseline/increased workload, standard/upgraded interfaces). This may be achieved either through overlapping Incident Response events onto other vignettes (e.g., Window of Opportunity, Contact Management), or through a standalone vignette involving several high priority tasks that must be prioritized based on crew discretion (safety, mission). The scenario events and consequent decision factors that should be considered for this vignette include:

- Fault Diagnosis – The identification or recognition of cues within the system to allocate a fault related to boat health and consequent safety implications to the crew. With respect to scenario simulation, representations of boat health indicators would occur on crew displays and have enough dynamic capability to generate changing states and corresponding alerts/cautions/warnings. System improvements in these areas would create opportunities to investigate performance impacts as they relate to areas of situation awareness and ability to respond to fault events;

- Threat Diagnosis – Similar to fault diagnoses of boat health, but involving additional aspects external to the boat. Within a scenario these situations can involve threats that arise from impending danger (e.g., counter-detection from enemy) or have already occurred (e.g., collision, torpedo launch), necessitating a reactive response to maintain the boat safety and well-being. In general, these events are of high priority and require immediate attention;
- Translation to Boat Capability – Interpretation of boat health status, particularly in the case of multiple degradations is fundamental to making correct decisions towards mitigating incidents. This refers to the overall understanding of the situation and any repercussions that each fault or threat may have on what the boat can (or cannot) achieve. The collective understanding of boat capability and any current degradations is essential towards attaining shared SA;
- Course of Action – The selected course(s) of action reflect the diagnosis and understanding of situations and provides an indication as to where priorities exist for the commander. Pending sufficient scenario development, researchers may potentially link these decisions to pass/fail outcomes;
- Conservation of Resources – Conservation during the execution of actions reflect the ability of the crew to be efficient while achieving their aim. During the interviews, SMEs recommended that battery conservation be considered as an indicator of performance effectiveness during experimentation;
- Situation Awareness – SA in the context of incident response is a crucial determinant of decision making and contributor towards successful coordination of crew collaboration. Several SMEs commented during interviews that boat health status displays can benefit from good human factors design, particularly if they can assist with decision making through improved information displays (e.g., integrated schematic visualizations). These types of interventions could be exercised during research to determine the impact of interface improvements as compared to baseline capability (existing displays).

4.4.4.2 Incident Response Basic Measures

The basic measures selected for the Incident Response vignette are presented below. These metrics were drawn from the Excel list and derived primarily from the Submarine Systems and SA task contexts:

- Diagnosis of fault (compared to ground truth) (OBJ, MOP) – Refers to basic state representations of boat health on the interfaces (e.g., air quality, temperature, engine status). Threshold limits (e.g., eroding margins of safety) could be used to trigger recognition events and consequent behaviours or priority tasks that follow. These may involve simple declarations, or scripted task sequences as determined by the experimental scenario;
- Battery and fuel remaining after incident (OBJ, MOP) – Dynamic representations of the battery and fuel consumption may involve layers of complexity outside the scope of research. As a simpler alternative, static state representations could be used to drive behavioural markers, declarations, or task sequences to evaluate performance. For example, starting a scenario with the battery at 25% may cause crews to favour execution of a snort before other tasks;

- Shared SA regarding boat health in relation to its ability to act (subject response to be evaluated by expert) (comprehension level 2 SA) (SUBJ, MOP) – Pertains to the interpretation of boat health (e.g., battery power) and the impact on her ability to carry out tasks (e.g., investigate contact at distance ‘D’, bearing ‘B’, heading ‘H’, at speed ‘S’);

4.4.4.3 Incident Response Composite Measures

The composite measures selected for the Incident Response vignette are presented below. These metrics were drawn from the Excel list and derived primarily from the Submarine Systems and SA task contexts:

- Reactive responses when encountering unexpected events (incidents) necessitating deviations from the existing plan (expert evaluation) (SUBJ, MOP) – Includes the assessment of course of action and re-planning activity in the event of inappropriate contingencies. Reactive scenarios can be achieved through partial omission of information or providing non-specific activity during a pre-experimental brief. Experimenters should be mindful to provide sufficient guidelines pertaining to mission conduct and priority goals so participants will respond within expectations;
- Communication of contact priorities to crew (expert evaluation) (SUBJ, MOP) – Ratings of behaviour that can be used to score participant actions when prioritizing tasks. The quantity and complexity of tasks will be driven by experimental objectives. This may involve tasks such as attending to boat health, tracking a COI, or evading a counter-detection; all of which can exist in different priority ordering depending on the situation being simulated;
- Was successful mission completion impacted by a system issue? (expert evaluation) (SUBJ, MOE) – A number of subjective measures can be tailored in the form of questions to investigate the impacts of compromised submarine systems through expert ratings. These measures would require a deliberate representation of boat health and how they are affected by simulation events, for example:
- What submarine systems are of primary concern? (subject response to be evaluated by expert) (SUBJ, MOP);
- What is the plan to accommodate the incident/fault and maintain safety, covertness and/or achieve the mission (subject response to be evaluated by expert) (SUBJ, MOP);
- What is the estimated time to fix the fault? (subject response to be evaluated by expert) (SUBJ, MOP);
- Was appropriate consideration given to submarine systems? (expert evaluation) (SUBJ, MOE).

4.4.4.4 IID Related Measures

There are a number of elements of the IID that could assist the incident response task. The difficulty is measuring the SME’s use of the elements. To this end, there are a number of standard MOPs and MOEs that should be applied to the IID.

- EMT Scan Patterns (OBJ, MOE) – the pattern by which the SME scans the IID to carry out tasks associated with incident response should be studied and compared with the scan

patterns (including verbal or audible information acquisition tasks) of current incident response tasks. Any differences should be noted, especially if the SME begins to omit information sources when using the IID. This latter observation may either be evidence of good design (i.e., that the information is presented or combined implicitly with other IID elements) or over-reliance on particular information sources. With familiarity scan patterns should become much more direct and less 'dense' with respect to the number of pauses.

- EMT Dwells (OBJ, MOP) – the time (percentage of time spent looking at IID, percentage of time overall) spent looking at specific IID frames should be studied and compared with the amount of time spent looking or attending to analogous information sources in the current control room. In particular, good design should result in fewer dwells of shorter duration, especially as familiarity with the system grows.
- Mouse Movements (OBJ, MOP) – similar to EMT Dwells above, this would measure the time (percentage of time spent interacting with IID, percentage of time overall) spent interacting with specific IID frames, and compared with the amount of time spent interacting with analogous information sources in the current control room. Good design should result in few interactions required for the SME to reach their goal, especially as familiarity with the system grows.
- IID Configuration (OBJ, MOE) – for window of opportunity purposes the IID would typically show the platform state in the lowest frame, or possibly the library or a tote (none of these views are currently defined). The views used by the SME, as well as what they switch to in which frame, should be tracked and studied. If the SME has to switch views frequently, this may indicate a poor design of the IID.

Additionally, there are a number of specific MOPs and MOEs that can be applied to the IID.

- Alerts (SUBJ, MOE) – the IID presents a number of configurable alerts. The alerts should result in better SA and, therefore, better decision making and problem solving for the SME. Specific SA probe questions should be developed and posed to the SME during an incident response vignette in both the IID and non-IID conditions. If the IID is well-designed, there should be fewer alerts or similar erosions of safety in the IID condition compared to the non-IID condition.
- The IID should also result in better SA regarding submarine systems, which could be evaluated through the use of probe questions. These measures would require a deliberate representation of boat health and how they are affected by simulation events, and should be repeated in both IID and non-IID conditions. Some example questions could include:
  - What submarine systems are of primary concern? (SUBJ, MOP);
  - What is the plan to accommodate the incident/fault and maintain safety, covertness and/or achieve the mission (SUBJ, MOP);
  - What is the estimated time to fix the fault? (SUBJ, MOP);
  - Was appropriate consideration given to submarine systems? (expert evaluation) (SUBJ, MOE).

Additionally, most of the basic and composite measures described above can be used in an IID/non-IID comparison to infer the impact of the IID.

5 Conclusions

The work described in this report addressed two objectives: the development of MOPs and MOEs for use in simulations and experiments of the VICTORIA Class Submarine control room and the repurposing of CWA and GDTA design work for MOP and MOE development. These objectives have been met with different degrees of success. The first objective has been fully met (see Section 4.4 and Annex A). Indeed, more MOPs and MOEs were developed for vVic than can realistically be applied. A selection of these has been presented for serious consideration in the main body of this report. The second objective, the use of CWA and GDTA products to develop the MOPs and MOEs, has only been partially successful (see Section 4.2). Much of the previous work seemed to achieve a type of convergent validation of itself, although it was encouraging that the GDTA products largely agreed with the CWA products. The functions and work situations represent strong guidance regarding the development of MOPs and MOEs to cover the range of the work domain. The analysis that came closest to addressing the depth of the work domain and developing MOPs and MOEs that stand on their own (i.e., they are ‘absolute’ in the sense that a clear distinction between success and failure can be made, irrespective of the specifics of the experimental scenario) was the Strategies Analysis. However, in common with other analyses done previously, the Strategies Analysis did not provide information that easily facilitated MOP and MOE development. In essence, the previous work did not ask the types of questions to facilitate MOP and MOE development. The previous work attempted a broad treatment of the VICTORIA Class Submarine work domain, rather than carrying out a detailed treatment of any one functional area.

While the previous work provided a good framework around which to develop MOPs and MOEs (specifically the functions identified), the previous analyses did not provide enough information to develop clear and unambiguous measures that would provide incisive and accurate evaluations of performance and effectiveness. This is not altogether unexpected, given what is known about complex, dynamic systems, of which Submarine Command and Control (C2) is an example. Interacting with other autonomous, thinking entities, in the open environment, always means that the range of possibly ‘correct’ actions, decisions, etc. is effectively infinite. Independent and absolute MOPs and MOEs are no doubt possible, but they are difficult to develop. Therefore subjective answers will be most expedient for the collection of performance and effectiveness data, and no unassailable evaluation can be made. The only way to ‘know’ what the right answer is, is to control the world of the submarine; that is, know about everything in that world, what it has done, and what it will do. In essence, it means controlling what is presented to the experiment participants and knowing beforehand what the ‘right’ or ‘best’ answer is. Because of this, we have recommended the development of a series of experimental vignettes that can be presented individually or together. When presented in one simulation, they offer a realistic range of activities that a typical submarine control room might expect to experience.

5.1 Summary of Experimental Vignettes

The following section presents a condensed summary of each vignette and recommended approach to application. In general these vignettes were contextually grouped so that they can be developed in isolation (individual tests) for future experimentation, or as phases of an extended mission scenario pending the level of complexity afforded by the vVic simulation environment:

- Navigation to Destination Vignette – Ship handling under normal circumstances will likely not yield a great deal of insight regarding performance. Common tactical challenges need to be introduced during ship handling, which includes the considerations of effective sensor and the overall maintenance of safety to better elicit observable performance. Much of this will be driven by scenario events and provide reason to exercise tactical movement, such as deviations from course to evade a contacts' counter-detection capability. This vignette requires representation of an interface capable of integrating some level of tactical symbology (e.g., own boat, known units) and navigation markers (e.g., boundaries, boat course). Section 4.4.1 provides a list of measures that can be used to evaluate performance in this context.

- Contact Management Vignette – The Contact Management vignette provides a high level of scalability based on its ability to adapt varying degrees of complexity, ranging from simple isolated COI scenarios to high density contact environments. The basic area of performance measurement within this vignette is assessing the operator's ability to collect information on a contact and narrowing down its threat level to determine the priority of consideration (urgency to act). Depending on the fidelity of the simulation, this may involve several tasks related to the application of TMA and the supporting surface or sub-surface sensor systems. This vignette is a good candidate to initiate comparative studies involving design interventions to assess whether performance can be improved through interface level modifications. In general, these can be determined through time based performance metrics such as the detection, identification, and classification of contacts.

- Window of Opportunity Vignette – The Window of Opportunity vignette encompasses the tasks related to surface (or near surface) activity that involve the maintenance of boat health (e.g., snorting), communications, and surface contacts. The proposed application of events within this vignette are to establish a foundation of work tasks representing 'opportunistic' events that may have been deliberately planned or routine duties that can be executed without significant urgency. This vignette will allow researchers to establish a baseline level of performance to characterize non-urgent performance and to assess the level of expediency in which a collective series of tasks can be completed (e.g., snort, communicate, and observe a surface contact within a surfacing window timeframe). Further investigations pertaining to workload impacts and response to urgency can be used to observe influence of greater task difficulty on performance. The candidate areas of interest would likely involve further multi-tasking and prioritization, as well as any design interventions that can be applied to resolve performance degradations resulting from increased task challenges.

- Incident Response Vignette – The Incident Response vignette is intended to evaluate reactive performance and re-planning tasks arising from higher priority events which take precedence over existing activity. The introduction of increased time pressure, urgency of task conduct, and re-prioritization will likely have significant impact on operator performance. It was suggested previously that the Incident Response vignette (4.4.4) be applied in a manner which allows researchers to compare against a baseline level of performance, where the effects of increased task demands can be observed, and where predictions of error or failure might occur. This level of insight, if attainable through experimentation, will be of great value to both designers of future systems (e.g., technological solutions), and the operational community as it applies to the training and approach in which to adopt new technology.

5.2 Summary of Recommendations

The recommendations gathered for this report are a compilation of the project insights developed following the CWA work product review, proposed MOP/MOE groupings, SME interview sessions, and the reduction of their feedback. These can be summarized as follows:

- Previous CWA products are not sufficient on their own to develop MOPs and MOEs – The availability of CWA work within this domain is beneficial as it provides a great deal of traceability to concepts pertaining to the design and analysis of the anticipated interface capability. This can be used as a framework to develop MOPs and MOEs, but is not sufficient. Rather, a parallel or staggered effort is required to develop MOPs and MOEs. This effort may not be as exhaustive as that of the CWA, but would ideally leverage the same opportunities to interact with the SMEs as the CWA;

- SME assisted scenario tailoring – Excessively scripted or controlled experimentation often leads to artificial and misrepresented task behaviour. SME consultation during scenario development significantly increases realism and further assists experimenters with determining where control is needed and where free flowing events are appropriate to gain better performance insights. In essence, this is the task of finding the optimal balance between internal validity (gain control by sacrificing scenario realism) and external validity (gain realism by sacrificing experimental control);

- Combine the objective with the subjective – The use of subjective metrics does not imply that performance cannot be quantified; nor does objective measurement imply an easy ‘black and white’ interpretation of results. The application of both types of measurement allows experimenters to mitigate the sacrifices made when balancing between internal and external validity (discussed previously). In effect, the combined application of both types of measurement will improve the end result, where objective metrics convey ‘what’ was observed (e.g., missed contact detection), and the subjective metrics help rationalize ‘why’ it occurred (e.g., higher perceived workload);

- Review training materials and procedures – Researchers should be cognizant of what is tested and how performance is evaluated within operational domains. This refers to the training paradigms used to acquire the fundamental skills needed to operate the VICTORIA Class Submarine. Such an understanding will allow researchers to develop a baseline for both the expected levels of ‘good’ performance, as well as the system capabilities (or limitations) that operators must work with to accomplish their tasks. Training materials and evaluation proforma were not reviewed as part of this contract;

- Connect scenario events to simulator system capability – It is necessary to provide a clear relationship between experimental measurements and the scenario factors which will stimulate the behaviours desired for observation. Most importantly, an accurate understanding of the interactive capability between the external environments (e.g., contact behaviours and attributes, dynamic versus static state representations) and the vVic submarine (e.g., interface fidelity and functionality) is key to managing expectations for future research;

- Influence the vVic development – Researchers have a vested interest in ensuring that capabilities match what they are looking to test. Any opportunities to be involved with defining the requirements for the vVic simulator should be exercised to better align the

supporting system capability with the desired research objectives. At minimum, researchers should establish enough interaction to understand what will be easier or harder to accomplish from the development perspective. This facilitates decisions on where to allocate effort during experimental planning;

- Select feasible experimental objectives – Select an experimental course of action or approach that is scalable to the capability available on the vVic simulation environment. The selection of these objectives will be made easier through an informed understanding of what is feasible (or not). The feasibility evaluation presented Section 4.3.2 of this report can be used as a guideline;

- Economy of effort should not always determine task priority – Expending project resources on the more difficult objectives should not be regarded as misdirected. It is often worthwhile to target investigations requiring greater effort as they may provide greater relevance to the operational community and consequently improved return on investment for R&D. A possible approach to tackling larger objectives is the development of several sub-goals through incremental stages of research. This allows researchers to build a succession of knowledge through a series of small projects and proceed with enough agility to determine whether further investigations will show promise;

- Cross functional scenarios are needed to evaluate performance – In order to contextually consider the functional work areas on board the VICTORIA Class Submarine, it is necessary to regard multiple factors within the environment. While the experimental vignettes are organized as distinct, potentially standalone settings in this report, researchers should consider the extension of these applications in concurrent or overlapping circumstances. This provides researchers the ability to significantly increase the complexity of task demands by challenging decision making, information management, and consequent degradations to workload and SA. For any project experimenting in performance and system design, it is of great value to be able to accurately reproduce areas of challenge and demonstrate whether interventions (e.g., improved interface displays or functions) will alleviate performance degradations;

- Of the scenario vignettes described in this work, the two most relevant vignettes are the navigation vignette and the contact management vignette. These two vignettes most intensively combine the overall goals of safety, covertness, and mission success. DRDC Atlantic should allocate significant time and effort to the development of a scenario in a challenging area of water (close to land, with many islands or sea mounts), with a mix of traffic (surface and air) as well as friendly and hostile forces. Such a scenario will challenge the navigation and contact management expertise of the VICTORIA Class control room;

- SME critiques of the (evolving) scenario are often very educational with regard the important factors in their job and can be used to develop additional MOPs and MOEs, as well as develop better scenarios; and,

- As far as possible, define all subjective data collection to the greatest possible extent. This will improve the speed of data collection, reduce the post-exercise analysis, and enhance inter-rater reliability.

5.3 Relationship to Future Work

The information contained within this report will support the future evolution of work related to the vVic simulator. The anticipated follow on work will be closely related to the following areas of work:

- Integrated Information Display – A substantial amount of design guidance and interface concepts will be drawn from the previous CWA reports (Taylor et al, 2009; Bruyn Martin et al, 2009; Bruyn Martin et al, 2010; Bruyn Martin & Taylor, 2010; Rehak et al, 2011a; Rehak et al, 2011b) in future designs. This report highlighted some areas of experimentation within the IID that could be pursed using the recommended scenario vignettes and MOP/MOEs (e.g., Section 4.4.2.4);
- Experimental Plan – It is anticipated that experimental planning will aim to asses operator performance within a framework of experimental scenarios and MOP/MOEs selected from this report. This will involve the vVic simulator and primarily utilize the IID functionality to establish performance based experimental designs and to determine the effectiveness of any design solutions identified; and,
- Simulator Development – It is expected that the researchers tasked with conducting experimentation on the IID will have regular interactions with the development team to understand the capability of the system and what can be accomplished within the forecasted project schedule and budget (if/when available). Ideally, researchers should aim to influence the development cycle through design recommendations that reflect the priority areas of research and the individual aspects of performance that they plan to investigate.

This page intentionally left blank.

6 References

[1] Bruyn Martin, L, Taylor, T. E., Karthaus, C., and Matthews, M. (2010). Unpublished contractor report.

[2] Bruyn Martin, L. and Taylor, T. E. (2010). Unpublished contract report.

[3] Bruyn Martin, L., Taylor, T., and Karthaus, C. (2009) Unpublished contract report

[4] Endsley, M.R., Bolte, B., and Jones, D.J. (2003). Designing for situation awareness: An approach to user-centered design. Laurence Erlbaum Associates, Mahwah, NJ.

[5] Hunter, A. & Hazen, M.G. (2011) Development of measures of effectiveness and performance for the command team aboard a VICTORIA class submarine. Human Research Ethics Protocol L-807. August 2011

[6] Lamoureux, T.M., Pronovost, S., and Dubreuil, A. (2011). Requirements for periscope simulation. DRDC Atlantic Contract Report, DRDC Atlantic CR 2011-091, August 2011.

[7] Rehak, L., Karthaus, C., Lee, B., Matthews, M., and Taylor, T. (2011a). Unpublished contract report

[8] Rehak, L., Karthaus, C., Lee, B., Matthews, M., and Taylor, T. (2011b). Unpublished contract report

[9] Taylor, T., Karthaus, C., and Bruyn Martin, L. (2009). Unpublished contract report.

[10] Vicente, K.J. (1999). Cognitive Work Analysis: Toward safe, productive and healthy computer-based work. Lawrence Erlbaum Associates, Mahwah, NJ.

This page intentionally left blank.

Annex A ANNEX A: Full List of MOPs and MOEs Developed

Metric	Difficulty / Evaluation	Type Metric	Task Context	Scope of Metric
Duration of communication	1	OBJ	comms	MOP
Individual elements of communication that are critical to successful exchange of information (quality)	2	SUBJ	comms	MOP
Responsiveness to information requests – immediate, delayed	2	SUBJ	comms	MOP
Was the mission successfully completed – yes or no?	3	OBJ	comms	MOE
Frequency of communications (with whom, scenario context, by what means)	3	OBJ	comms	MOP
Intercepted communication (EW)	3	OBJ	comms	MOP
Degree to which communications external to the VCS can be delivered efficiently and effectively (overall rating scale)	4	SUBJ	comms	MOE
Appropriateness – timely, necessary, succinct, sufficient	4	SUBJ	comms	MOE
Command team was in possession of all necessary information	4	SUBJ	comms	MOE
Command team shared all necessary information	4	SUBJ	comms	MOE
Detection of anamoly traffic (contact behaviour) for within mission scenarios	1	OBJ	contact mgmt	MOP
Threshold: Angle on the Bow assessed within 5 - 10 deg accuracy	1	OBJ	contact mgmt	MOP
Threshold: Range estimation within 10% accuracy	1	OBJ	contact mgmt	MOP
Speed estimation based on range bearing readings (at least 2)	1	OBJ	contact mgmt	MOP
Was an accurate tactical picture maintained throughout the scenario (composite percentage of contacts detected/tracked and ID level)	1	OBJ	contact mgmt	MOE
Number of 'lost' contact incidences	1	OBJ	contact mgmt	MOE
Duration of time contacts were positively tracked vs duration of time contacts were out of contact	1	OBJ	contact mgmt	MOE
Number of contacts detected vs number in scenario	1	OBJ	contact mgmt	MOP
Number of 'Unknowns'	1	OBJ	contact mgmt	MOP
Knowledge of current and future positions of all contacts in scenario	1	OBJ	contact mgmt	MOP
Number of COIs found and tracked (percentage)	1	OBJ	contact	MOP

Metric	Difficulty / Evaluation	Type Metric	Task Context	Scope of Metric
			mgmt	
Contact declarations to narrow threat class	2	SUBJ	contact mgmt	MOP
Appropriate selection of video size / quality	2	SUBJ	contact mgmt	
Contact priorities (divide into first, second, third/not a priority) (subject response to be evaluated by expert)	2	SUBJ	contact mgmt	MOP
Communication of contact priorities to crew (expert evaluation)	2	SUBJ	contact mgmt	MOP
Multi-tasking: maximize information gathering within a time window	3	OBJ	contact mgmt	MOE
Contact re-classification, false alarms, or repeated contacts	3	OBJ	contact mgmt	MOP
Was the mission successfully completed – yes or no?	3	OBJ	contact mgmt	MOE
Amplification level achieved for each/all contacts	3	OBJ	contact mgmt	MOP
Accuracy of TMA when compared to ground truth (course, bearing, range, speed)	3	OBJ	contact mgmt	MOP
Speed and accuracy of go deep/look interval calculations	3	OBJ	contact mgmt	MOP
Number of contacts that an officer can track for the purposes of look interval	3	OBJ	contact mgmt	MOP
Accuracy of predictions of collision threats	3	OBJ	contact mgmt	MOP
Completeness of tactical picture (expert evaluation)	4	SUBJ	contact mgmt	MOE
Confidence in understanding of tactical picture (self-report)	4	SUBJ	contact mgmt	MOE
Problem solving tradeoff between tracking contacts and dropping contacts (expert evaluation)	4	SUBJ	contact mgmt	MOE
Clarity of contact plot (i.e., no unnecessary stale, lost, time late contacts) (expert evaluation)	4	SUBJ	contact mgmt	MOE
Workload questionnaire/NASA TLX (self-report)	4	SUBJ	contact mgmt	MOE
Overemphasis on any element of the a priori information, intel or ORBAT (expert evaluation)	4	SUBJ	contact mgmt	MOP
Effectiveness of sensor employment (e.g., periscope vs towed array vs onboard sonar (active/passive) vs comms vs intel vs radar vs EW) (expert evaluation)	4	SUBJ	contact mgmt	MOP
Time spent at periscope depth	1	OBJ	covertness	MOP
Successful evasion	3	OBJ	covertness	MOE
Number of counter-detections	3	OBJ	covertness	MOE
Was the mission successfully accomplished - yes or no?	3	OBJ	covertness	MOE

Metric	Difficulty / Evaluation	Type Metric	Task Context	Scope of Metric
Time spent communicating	3	OBJ	covertness	MOP
Frequency / duration of non-covert actions (active sonar, comms, etc.)	3	OBJ	covertness	MOP
Cavitations	3	OBJ	covertness	MOP
An appropriate balance of communication (or other detectable actions) was maintained (rating scale)	4	SUBJ	covertness	MOE
Adequacy of decision making regarding covertness trade-offs	4	SUBJ	covertness	MOE
Appropriate management of own ship signature	4	SUBJ	covertness	MOP
Management of noise level in the control room and submarine more generally (noise husbandry)	4	SUBJ	covertness	MOP
Appropriate selection of environment (noise, traffic, temperature, bottom topography)	4	SUBJ	covertness	MOP
Range and bearing to specified contacts (perception level 1 SA)	1	OBJ	individual SA	MOP
Time remaining before next look (perception level 1 SA)	1	OBJ	individual SA	MOP
Battery status (perception level 1 SA)	1	OBJ	individual SA	MOP
Speed, depth, course of submarine (perception level 1 SA)	1	OBJ	individual SA	MOP
Top 3 priority contacts and why (self-report) (comprehension level 2 SA)	2	SUBJ	individual SA	MOP
Likely actions of top 3 priority contacts in next 10 min, 30 min, 1 hr (self-report) (projection level 3 SA)	2	SUBJ	individual SA	MOP
Mission priority at that time (safety, covertness or a specific mission objective) (self-report) (comprehension level 2 SA)	2	SUBJ	individual SA	MOP
Groups of contacts and why are they grouped (self-report) (comprehension level 2 SA)	2	SUBJ	individual SA	MOP
Was the mission successfully completed – yes or no? (comprehension level 2 SA)	3	OBJ	individual SA	MOE
Sea state (perception level 1 SA)	3	OBJ	individual SA	MOP
Winds (perception level 1 SA)	3	OBJ	individual SA	MOP
Bearing of sun/moon (perception level 1 SA)	3	OBJ	individual SA	MOP
Confidence that the submarine system status (now and in the near future) was understood (self-report) (comprehension and projection, level 2 & 3 SA)	4	SUBJ	individual SA	MOE
Confidence that the tactical picture (now and in the near future) was understood (self-report) (comprehension and projection, level 2 & 3 SA)	4	SUBJ	individual SA	MOE

Metric	Difficulty / Evaluation	Type Metric	Task Context	Scope of Metric
Confidence that the mission would be achieved with no compromise to safety or covertness (self-report) (comprehension and projection, level 2 & 3 SA)	4	SUBJ	individual SA	MOE
Degree of understanding of the tactical picture (subject response to be evaluated by expert) (comprehension and projection, level 2 & 3 SA)	4	SUBJ	individual SA	MOE
Degree of understanding of submarine system status (subject response to be evaluated by expert) (comprehension and projection, level 2 & 3 SA)	4	SUBJ	individual SA	MOE
Degree to which mission objectives were being satisfied, including safety and covertness (expert evaluation) (comprehension and project, level 2 & 3 SA)	4	SUBJ	individual SA	MOE
Number and duration of 'dwells' if using eye movement tracking (perception level 1 SA)	5	OBJ	individual SA	MOP
Accuracy of synchronization with other assets	1	OBJ	planning	MOE
Completeness of information contained in navigation plan (e.g., speed, heading, leg duration, depth, snorting interval, look interval, communications, engineering, life support, domestics, sensor use, weapons use)	2	SUBJ	planning	MOP
Clarity of definition of safety thresholds and success criteria (expert evaluator)	2	SUBJ	planning	MOP
Was the mission successfully completed – yes or no?	3	OBJ	planning	MOE
Duration of actual mission vs planned duration	3	OBJ	planning	MOE
Number of mission objectives achieved (percentage)	3	OBJ	planning	MOE
Number of identified mission milestones achieved (percentage e.g., comms, waypoints, rendezvous, etc.)	3	OBJ	planning	MOE
Time spent involved in planning activities	3	OBJ	planning	MOP
Number of COAs developed	3	OBJ	planning	MOP
Number of contingencies (branches, sequels) developed	3	OBJ	planning	MOP
Number of effects brought to bear in mission	3	OBJ	planning	MOP
Duration of planned mission	3	OBJ	planning	MOP
Reactive responses when things don't go according to plan	4	SUBJ	planning	MOP
Degree to which plan was followed to successfully achieve mission (expert evaluator)	4	SUBJ	planning	MOE
Degree to which any deviations were accounted for in contingency planning (expert evaluator)	4	SUBJ	planning	MOE
Effective employment of effects (expert evaluator)	4	SUBJ	planning	MOE
Efficiency of crew in executing plan (expert evaluator)	4	SUBJ	planning	MOE
Coordination of crew in executing plan (expert evaluator)	4	SUBJ	planning	MOE

Metric	Difficulty / Evaluation	Type Metric	Task Context	Scope of Metric
'Elegance' of plan (i.e., smooth transition from one milestone in the plan to the next, economical) (expert evaluator)	4	SUBJ	planning	MOP
Adequacy of consideration of safety, covertness, mission (expert evaluator)	4	SUBJ	planning	MOP
Adequacy of red team and wargaming (expert evaluator)	4	SUBJ	planning	MOP
Adequacy of consideration of prior information, commander's intent, higher orders, etc. (expert evaluator)	4	SUBJ	planning	MOP
Effectiveness of plan communication to crew and higher/TG (expert evaluator)	4	SUBJ	planning	MOP
Collision with vessels or land	1	OBJ	safety	MOE
Closest Point of Approach (CPA) (Range, Bearing)	1	OBJ	safety	MOP
Look interval (duration)	1	OBJ	safety	MOP
Frequency of going deep (safe depth)	1	OBJ	safety	MOP
Frequency and interval of all round looks	1	OBJ	safety	MOP
Number of Milestone ARLs (All Round Look)	1	OBJ	safety	MOP
Bearing rate tolerance (within 'x' meters per second)	3	OBJ	safety	MOP
Was the mission successfully completed – yes or no?	3	OBJ	safety	MOE
Were any submarine sub-systems operated at a degraded state – yes or no?	3	OBJ	safety	MOE
Health indicators – fuel, engine status, air (eroding margin of safety)	3	OBJ	safety	MOP
Pilotage (speed, bank, pitch, roll, depth) as they approach threshold	3	OBJ	safety	MOP
Shared SA regarding Operational goals	4	SUBJ	safety	MOE
Distance into the future that situational expectations extend (time)	4	SUBJ	safety	MOE
Confidence that crew had a strong mutual understanding of goals and how to achieve them	4	SUBJ	safety	MOE
Frequency of human errors – occurrence of omission and commission (extra action, wrong action, wrong order) (rating)	4	SUBJ	safety	MOP
Workload questionnaire or NASA-TLX	4	SUBJ	safety	MOP
Appropriateness of safe depth	4	SUBJ	safety	MOP
Shared SA regarding priority contacts (and why) (subject response to be evaluated by expert) (comprehension level 2 SA)	2	SUBJ	shared SA	MOP
Was the mission successfully completed – yes or no? (comprehension level 2 SA)	3	OBJ	shared SA	MOE
Frequency of requests for clarification (comprehension level 2 SA)	3	OBJ	shared SA	MOP
Frequency of pre-emptive behaviour (point-outs, "would you like me to…") (comprehension and	3	OBJ	shared SA	MOP

Metric	Difficulty / Evaluation	Type Metric	Task Context	Scope of Metric
projection level 2 & 3 SA)				
Frequency of updates pertaining to submarine environment; future plans, knowledge of mission progress, crew status, internal sub-status, sub location (comprehension and projection level 2 & 3 SA)	3	OBJ	shared SA	MOP
Frequency of corrective commands by the OOW or Watch Leader or CO (comprehension level 2 SA)	3	OBJ	shared SA	MOP
Shared SA regarding Operational goals (subject response to be evaluated by expert) (comprehension and projection level 2 & 3 SA)	4	SUBJ	shared SA	MOE
Distance into the future that situational expectations extend (time) (self-report) (projection level 3 SA)	4	SUBJ	shared SA	MOE
Confidence that crew had a strong mutual understanding of goals and how to achieve them (self-report) (comprehension and projection level 2 & 3 SA)	4	SUBJ	shared SA	MOE
Teamwork questionnaire (self-report)	4	SUBJ	shared SA	MOE
Shared SA regarding boat health (e.g., battery power) and status (bearing to contact, heading) (subject response to be evaluated by expert) (comprehension level 2 SA)	4	SUBJ	shared SA	MOP
Shared SA regarding Operational goals (subject response to be evaluated by expert) (comprehension and projection level 2 & 3 SA)	4	SUBJ	shared SA	MOP
Shared SA regarding Commander's intent (subject response to be evaluated by expert) (comprehension and projection level 2 & 3 SA)	4	SUBJ	shared SA	MOP
Workload of the control room operators (self-report) (comprehension level 2 SA)	4	SUBJ	shared SA	MOP
Challenges or support requirements of the control room operators (self-report) (comprehension and projection level 2 & 3 SA)	4	SUBJ	shared SA	MOP
Nature of the request for clarification (subject response to be evaluated by expert) (comprehension level 2 SA)	4	SUBJ	shared SA	MOP
Trust questionnaire (self-report)	4	SUBJ	shared SA	MOP
Look interval (duration)	1	OBJ	ship handle	MOP
Appropriate checks of depth and speed before raising mast	2	SUBJ	ship handle	MOP
Adherence to boxes (friendly boundaries, routing, etc.)	3	OBJ	ship handle	MOP
% of time spent blind in certain areas	3	OBJ	ship handle	MOP

Metric	Difficulty / Evaluation	Type Metric	Task Context	Scope of Metric
% time spent observing mission critical areas	3	OBJ	ship handle	MOP
Fuel burn	3	OBJ	ship handle	MOE
Time submerged	3	OBJ	ship handle	MOE
Comparison of submarine historical track with the planned (deviation)	3	OBJ	ship handle	MOE
Cavitations	3	OBJ	ship handle	MOP
Surface and descent speed (rate)	3	OBJ	ship handle	MOP
Omission of required orders	3	OBJ	ship handle	MOP
Time taken to calculate angle on the bow (AOB)	3	OBJ	ship handle	MOP
Time taken to do bearing rate calculations	3	OBJ	ship handle	MOP
Correct or appropriate actions are taken in response to current tactical situation	4	SUBJ	ship handle	MOE
Successful execution of navigation plan	4	SUBJ	ship handle	MOE
Maintenance of submarine position in contacts baffles	4	SUBJ	ship handle	MOE
Degree to which elements of seamanship are successfully executed	4	SUBJ	ship handle	MOP
Maintenance of appropriate boat depth	4	SUBJ	ship handle	MOP
Appropriately putting the periscope on the next required bearing	4	SUBJ	ship handle	MOP
Proportion of time engaged in planning behaviours (% estimate or average per hour)	4	SUBJ	ship handle	MOP
Diagnosis of fault provided schematic interface diagrams	3	OBJ	sub systems	MOP
Successful maintenance of atmospheric conditions	3	OBJ	sub systems	MOE
Was the mission successfully completed – yes or no?	3	OBJ	sub systems	MOE
Fuel remaining	3	OBJ	sub systems	MOE
Battery level remaining	3	OBJ	sub systems	MOE
Air quality	3	OBJ	sub systems	MOE
Comparison of time estimate with actual time to fix/make safe	3	OBJ	sub systems	MOP

Metric	Difficulty / Evaluation	Type Metric	Task Context	Scope of Metric
Diagnosis of fault (compared to ground truth)	3	OBJ	sub systems	MOP
Time to next snort	3	OBJ	sub systems	MOP
Air quality (oxygen, CO2, hydrogen, etc.)	3	OBJ	sub systems	MOP
Battery remaining	3	OBJ	sub systems	MOP
Diesel remaining	3	OBJ	sub systems	MOP
Was successful mission completion impacted by a system issue? (expert evaluation)	4	SUBJ	sub systems	MOE
Was appropriate consideration given to submarine systems? (expert evaluation)	4	SUBJ	sub systems	MOE
Was safety ever threatened by a submarine system fault? (subject response to be evaluated by expert)	4	SUBJ	sub systems	MOE
Was covertness ever threatened by a submarine system fault? (subject response to be evaluated by expert)	4	SUBJ	sub systems	MOE
What submarine systems are of primary concern? Why? (subject response to be evaluated by expert)	4	SUBJ	sub systems	MOP
What is the estimated time to fix the fault? (subject response to be evaluated by expert)	4	SUBJ	sub systems	MOP
What is the plan to accommodate the incident/fault and maintain safety, covertness and/or achieve the mission (subject response to be evaluated by expert)	4	SUBJ	sub systems	MOP

Annex B ANNEX B: PowerPoint Slides Used for SME Validation

B.1 SME Validation Slides

Agenda

- Brief introduction and objectives
- Scoring approach
- MOPs and MOEs

Introduction

- Several pieces of work done to analyze the VICTORIA Class work domain
  - Specifically Control Room and Command Team
- Work culminated with design for Integrated Information Display
- Also, virtual VICTORIA (vVic) has been built at DRDC Atlantic
  - Full-size mock up of Control Room
  - To be used for a variety of purposes, including experimentation

Objectives

- Objective of this work twofold:
  - Determine whether previous analyses can be repurposed for development of MOPs and MOEs
  - Development of MOPs and MOEs for vVic
- MOPs: lower level metrics of task performance
  - Typically assess things an individual would do
  - E.g. steer true to course, track a contract, detect a signal
- MOEs: higher level metrics of system effectiveness
  - Typically assess things the whole system (vehicle, sensors, weapons, humans) works to achieve
  - E.g. achieve mission, manage contacts, maintain safety, operate as a member of the task group

Scoring Approach

- Each slide presents one measurement ‘area’
- Measurement area is described according to the ‘essence’ of the function that we would want to measure
  - You rate your agreement on a scale of 1 (strongly disagree) to 5 (strongly agree)
  - Suggest any alternative descriptions; we will note these
- MOPs described
  - Split into objective and subjective
  - You rate the ‘appropriateness’ (is it valid, is it measurable) of each MOP on a scale of 1 – 5
  - Suggest any alternative MOPs; we will note these
- Repeat for MOEs for that measurement area

Summary of Discussion Areas

- Communications (external)
- Shared SA (includes internal crew interaction)
- Safety
- Ship Handling
- Covertness
- Planning
- Contact Management
- Individual SA
- Submarine Systems

General Mission Overview

Given the broad nature of some of measurement areas to be discussed, it is necessary to provide some guiding context

- Covert Tracking Scenario
- Submarine has been tasked to conduct covert surveillance and tracking of maritime traffic the southwest approaches to the English Channel.
- The mission is to obtain information on the normal activity in the area and identify any anomalus behaviour.
- The area of operations is a 25 x 40 nmi box south of the port of Penzance starting 2nmi off shore, other NATO units will be operating in adjacent water areas.
- Unit will maintain normal communications cycles.
- Local units may or may not be operating AIS in accordance with IMO regulations and space based radar contact maps are available twice daily.
- Unit is authorized to enter national waters if required but remaining covert is deemed essiential.
- National maritime air assets are shore and ship launched helo. No known MPA assets in the area. Naval units may be operating in the area.
- Water depths 50 - 150 m, shallow surface duct (15m). Weather is overcast, wind from south 15kts, SS 3-4. Forecast is to stay the same. Tides are 2-4 m, High Tide at 0900
- Expect Heavy shipping in Channel, Light shipping into Penzance. Heavy local fishing. Light pleasure boat traffic.
- Permission to use weapons in self-defence.

Communications

<u>Description</u>:

Refers primarily to the **external** communications, such as those involving the HHQ, TG, SUBOPAUTH. Any type of external communication should be considered, including those needed to acquire or exchange information related to tactical operations, and various sources of Intelligence. Security and covertness may be considered here if there are relevant options to “perform” these communications more or less effectively.

<u>MOPs</u>

Objective

- Frequency of communications (with whom, scenario context, by what means)
- Duration of communication
- Intercepted communication

Subjective

- Individual elements of communication that are critical to successful exchange of information (expert evaluator)
- Appropriateness – timely, necessary, succinct, sufficient (expert evaluator)
- Responsiveness to information requests – immediate, delayed (expert evaluator)

<u>MOEs</u>

Objective

- Was the mission successfully completed – yes or no?

Subjective

- Degree to which communications external to the VCS can be delivered efficiently and effectively (expert evaluator)
- Appropriateness – timely, necessary, succinct, sufficient (expert evaluator)
- Command team in possession of all necessary information (expert evaluator)
- Command team shared all necessary information (expert evaluator)

Situation Awareness (SA) – Preface

Discussion of SA will require a “level’ of consideration within subsequent slides; examples are provided in these later discussions. The general framework will be:

- Perception (Level 1)
  - What is the current level of…
  - Has … increased/decreased in the last two minutes
  - How long has the … been doing something
  - What actions have you taken?
  - What does the contact typically do?
- Comprehension (Level 2)
  - Is the contact moving/acting normally?
  - Are the contacts movements/actions consistent with their assumed state of mind?
  - Do you need to do anything else with this contact at this point in time?
- Projection (Level 3)
  - Do you expect the contact’s … to increase/decrease/change in the next minute or so?
  - Do you expect any changes in the contact in the next minute?
  - How long do you think you have until…?

Shared Situation Awareness (SA)

Description:

The maintenance or development of SA related to own ship awareness. Focal areas of discussion should involve the sharing, updating, and communication of information among the OOW or Watch Leader and crew members. SA area of interest may include Shared Command Intent (Knowledge of Mission Progress, Future Plans), Submarine Environment (Incident Management, Sub-status, Sub-location).

MOPs

Objective

- Frequency of requests for clarification (comprehension level 2 SA)
- Frequency of pre-emptive behaviour (point-outs, "would you like me to…") (comprehension and projection level 2 & 3 SA)
- Frequency of updates pertaining to submarine environment; future plans, knowledge of mission progress, crew status, internal sub-status, sub location (comprehension and projection level 2 & 3 SA)
- Frequency of corrective commands by the OOW or Watch Leader or CO (comprehension level 2 SA)

Subjective

- Shared SA regarding boat health (e.g., battery power) and status (bearing to contact, heading) (subject response to be evaluated by expert) (comprehension level 2 SA)
- Shared SA regarding priority contacts (and why) (subject response to be evaluated by expert) (comprehension level 2 SA)
- Shared SA regarding Operational goals (subject response to be evaluated by expert) (comprehension and projection level 2 & 3 SA)
- Shared SA regarding Commander's intent (subject response to be evaluated by expert) (comprehension and projection level 2 & 3 SA)
- Workload of the control room operators (self report) (comprehension level 2 SA)
- Challenges or support requirements of the control room operators (self report) (comprehension and projection level 2 & 3 SA)
- Nature of the request for clarification (subject response to be evaluated by expert) (comprehension level 2 SA)
- Trust questionnaire (self report)

Shared Situation Awareness (SA)

Description:

The maintenance or development of SA related to own ship awareness. Focal areas of discussion should involve the sharing, updating, and communication of information among the OOW or Watch Leader and crew members. SA area of interest may include Shared Command Intent (Knowledge of Mission Progress, Future Plans), Submarine Environment (Incident Management, Sub-status, Sub-location).

MOEs

Objective

- Was the mission successfully completed – yes or no? (comprehension level 2 SA)

Subjective

- Shared SA regarding Operational goals (subject response to be evaluated by expert) (comprehension and projection level 2 & 3 SA)
- Distance into the future that situational expectations extend (time) (self report) (projection level 3 SA)
- Confidence that crew had a strong mutual understanding of goals and how to achieve them (self report) (comprehension and projection level 2 & 3 SA)
- Teamwork questionnaire (self report)

Safety

<u>Description</u>:

A broad overarching goal for the submarine, generally referring to ensuring the safety of the crew, the integrity of the submarine itself, and the goal of keeping damage to other vessels or people at a minimum. This category can be regarded as a high level goal comprised of sub-components involving aspects such as Seamanship, Covertness, and understanding Submarine Systems.

<u>MOPs</u>

Objective

- Closest Point of Approach (CPA) (Range, Bearing)
- Health indicators – fuel, engine status, air (eroding margin of safety)
- Pilotage (speed, bank, pitch, roll, depth) as they approach threshold
- Look interval (duration)
- Frequency of going deep (safe depth)
- Frequency and interval of all round looks
- Number of Milestone ARLs (All Round Look)

Subjective

- Frequency of human errors – occurrence of omission and comission (extra action, wrong action, wrong order) (self report)
- Workload questionnaire or NASA-TLX (self report)
- Appropriateness of safe depth (subject response to be evaluated by expert)

Safety

<u>MOEs</u>

Objective

- Was the mission aborted for safety reasons – yes or no?
- Were any submarine sub-systems operated at a degraded state – yes or no?
- Collision with vessels or land

Subjective

- Is the ship operating in a manner where the ship is less likely to incur damage? (expert evaluator)
- Was the submarine operated safely (expert evaluator)
- Degree to which orders from higher have been met or exceeded (e.g., getting in closer traded off with risk tolerance, confidence in crew) (expert evaluator)

Ship Handling

Description:

Refers to the skilled maneuvering of the ship during various boat states or to maintain appropriate course with respect to the navigation plan. Includes the considerations and understanding of the submarine tactical disposition as they pertain to ship control, speed, and course in response to mission goals and/or response to incidents. Prioritization of the factors influencing seamanship and conditional courses of action may be discussed here.

MOPs

Objective

- Cavitation
- Surface and descent speed (rate)
- Omission of required orders
- Time taken to calculate angle on the bow (ATB)
- Time taken to do bearing rate calculations
- Look interval (duration)

Subjective

- Degree to which elements of seamanship are successfully executed (expert evaluator)
- Maintenance of appropriate boat depth (expert evaluator)
- Appropriately putting the periscope on the next required bearing (expert evaluator)
- Appropriate checks of depth and speed before raising mast (expert evaluator)
- Proportion of time engaged in planning behaviours (% estimate or average per hour) (expert evaluator)

Ship Handling

MOEs

Objective

- Fuel burn
- Time submerged
- Comparison of submarine historical track with the planned (deviation)

Subjective

- Correct or appropriate actions are taken in response to current tactical situation (expert evaluator)
- Successful execution of navigation plan (expert evaluator)

Covertness

Description:

Refers to the general goal of remaining undetected by hostile and neutral forces. Discussion of specific work actions such as signature management, evasion, surveillance, reconnaissance, and the communications-covertness tradeoff are key to identifying metrics of performance and overall effectiveness.

MOPs

Objective

- Time spent at periscope depth
- Time spent communicating
- Frequency / duration of non-covert actions (active sonar, comms, etc.)
- Cavitation

Subjective

- Appropriate management of own ship signature (expert evaluator)
- Management of noise level in the control room and submarine more generally (noise husbandry) (expert evaluator)
- Appropriate selection of environment (noise, traffic, temperature, bottom topography)(expert evaluator)
- Maintenance of submarine position in contacts baffles (expert evaluator)

Covertness

MOEs

Objective

- Successful evasion
- Number of counter-detections
- Was the mission successfully accomplished – yes or no?

Subjective

- An appropriate balance of communication (or other detectable actions) was maintained (rating scale) (expert evaluator)
- Adequacy of decision making regarding covertness trade-offs (expert evaluator)

Planning

<u>Description</u>:

The high level and detailed preparation carried out in advance of an operation, transit, activity, etc. Planning includes the collection of relevant data, including orders from higher commands, the consideration of the data, and the development of detailed courses of action, including contingency planning. Planning also includes setting safety thresholds, and mission success criteria. Subsequent to the planning activities carried out beforehand is the monitoring of an operation, transit, activity, etc. against the plan and the decision making and problem solving associated with enacting a contingency plan or reacting in an ad-hoc manner.

<u>MOPs</u>

Objective

- Time spent involved in planning activities
- Number of COAs developed
- Number of contingencies (branches, sequels) developed
- Completeness of information contained in navigation plan (e.g. speed, heading, leg duration, depth, snorting interval, look interval, communications, engineering, life support, domestics, sensor use, weapons use)
- Number of effects brought to bear in mission
- Duration of planned mission

Subjective

- ‘Elegance’ of plan (i.e. smooth transition from one milestone in the plan to the next, economical) (expert evaluator)
- Adequacy of consideration of safety, covertness, mission (expert evaluator)
- Adequacy of red team and wargaming (expert evaluator)
- Adequacy of consideration of prior information, commander’s intent, higher orders, etc. (expert evaluator)
- Clarity of definition of safety thresholds and success criteria (expert evaluator)
- Effectiveness of plan communication to crew and higher/TG (expert evaluator)

Planning

<u>MOEs</u>

Objective

- Was the mission successfully completed – yes or no?
- Duration of actual mission vs planned duration
- Number of mission objectives achieved (percentage)
- Number of identified mission milestones achieved (percentage e.g. comms, waypoints, rendezvous, etc.)
- Accuracy of synchronization with other assets

Subjective

- Degree to which plan was followed to successfully achieve mission (expert evaluator)
- Degree to which any deviations were accounted for in contingency planning (expert evaluator)
- Effective employment of effects (expert evaluator)
- Efficiency of crew in executing plan (expert evaluator)
- Coordination of crew in executing plan (expert evaluator)

Contact Management

<u>Description</u>:

The process of contact management, from initial detection, identification and tracking, through to prosecution of the contact (if required). Includes correlation of detection information with intelligence data, as well as the effective use of sensors to gather more information on the contact for correlation with a priori data (e.g. ORBAT). Also includes decision making regarding collision threats, go deep levels, look intervals, and the employment of additional assets to maximize effectiveness. Includes the effective problem solving (e.g. TMA) for tracking and targeting a contact, and the effective use of weapons or shows of force to achieve mission objectives.

<u>MOPs</u>

Objective

- Number of contacts detected vs number in scenario
- Amplification level achieved for each/all contacts
- Number of ‘Unknowns’
- Accuracy of TMA when compared to ground truth (course, bearing, range, speed)
- Speed and accuracy of go deep/look interval calculations
- Number of contacts that an officer can track for the purposes of look interval
- Accuracy of predictions of collision threats
- Knowledge of current and future positions of all contacts in scenario
- Number of COIs found and tracked (percentage)

Subjective

- Contact priorities (divide into first, second, third/not a priority) (subject response to be evaluated by expert)
- Overemphasis on any element of the a priori information, intel or ORBAT (expert evaluation)
- Effectiveness of sensor employment (e.g. periscope vs towed array vs onboard sonar (active/passive) vs comms vs intel vs radar vs EW) (expert evaluation)
- Communication of contact priorities to crew (expert evaluation)

Contact Management

<u>MOEs</u>

Objective

- Was the mission successfully completed – yes or no?
- Was an accurate tactical picture maintained throughout the scenario (composite percentage of contacts detected/tracked and ID level)
- Number of ‘lost’ contact incidences
- Duration of time contacts were positively tracked vs duration of time contacts were out of contact

Subjective

- Completeness of tactical picture (expert evaluation)
- Confidence in understanding of tactical picture (self report)
- Problem solving tradeoff between tracking contacts and dropping contacts (expert evaluation)
- Clarity of contact plot (i.e. no unnecessary stale, lost, time late contacts) (expert evaluation)
- Workload questionnaire/NASA TLX (self report)

Individual SA

Description:

The ability of the individual to be sensitive to changes in the environment in order that they understand the current situation and can make accurate predictions of how the situation will evolve in the near to medium future. 'Situation' includes the submarine's systems (e.g. fuel, propulsion, batteries, air, water, food, etc.), the morale and capabilities of the crew, the tactical situation (e.g. other vessels, contacts, aircraft, etc. in the area of interest), the location, depth and disposition of the submarine, the environmental conditions, and any other factors that affect the individual's decision making.

MOPs

Objective

- Probe questions: *"What is the/your…?"*
- Range and bearing to specified contacts (perception level 1 SA)
- Time remaining before next look (perception level 1 SA)
- Battery status (perception level 1 SA)
- Speed, depth, course of submarine (perception level 1 SA)
- Sea state (perception level 1 SA)
- Winds (perception level 1 SA)
- Bearing of sun/moon (perception level 1 SA)
- Number and duration of 'dwells' if using eye movement tracking (perception level 1 SA)

Subjective

- Probe questions: *"what are the…?"*
- Top 3 priority contacts and why (self report) (comprehension level 2 SA)
- Likely actions of top 3 priority contacts in next 10 min, 30 min, 1 hr (self report) (projection level 3 SA)
- Mission priority at that time (safety, covertness or a specific mission objective) (self report) (comprehension level 2 SA)
- Groups of contacts and why are they grouped (self report) (comprehension level 2 SA)

Individual SA

MOEs

Objective

- Was the mission successfully completed – yes or no? (comprehension level 2 SA)

Subjective

- Confidence that the submarine system status (now and in the near future) was understood (self report) (comprehension and projection, level 2 & 3 SA)
- Confidence that the tactical picture (now and in the near future) was understood (self report) (comprehension and projection, level 2 & 3 SA)
- Confidence that the mission would be achieved with no compromise to safety or covertness (self report) (comprehension and projection, level 2 & 3 SA)
- Degree of understanding of the tactical picture (subject response to be evaluated by expert) (comprehension and projection, level 2 & 3 SA)
- Degree of understanding of submarine system status (subject response to be evaluated by expert) (comprehension and projection, level 2 & 3 SA)
- Degree to which mission objectives were being satisfied, including safety and covertness (expert evaluation) (comprehension and project, level 2 & 3 SA)

Submarine Systems

Description:

Submarine systems refers to the level of knowledge and understanding required to safely and stealthily complete the assigned mission successfully. This includes management of systems entering and used by the control room (e.g. sensors, weapons, propulsion, steering, pitch, roll, etc.), atmospheric/environmental control, and power. It also includes incident management, understanding how long the incident will require to be made safe or rectified, knowing the moment-by-moment status of the incident, and understanding/predicting the impact the incident will have on mission goals.

MOPs

Objective

- Comparison of time estimate with actual time to fix/make safe
- Diagnosis of fault (compared to ground truth)
- Time to next snort
- Air quality (oxygen, CO_2, hydrogen, etc.)
- Battery remaining
- Diesel remaining

Subjective

- What submarine systems are of primary concern? Why? (subject response to be evaluated by expert)
- What is the estimated time to fix the fault? (subject response to be evaluated by expert)
- What is the plan to accommodate the incident/fault and maintain safety, covertness and/or achieve the mission (subject response to be evaluated by expert)

Submarine Systems

MOEs

Objective

- Was the mission successfully completed – yes or no?
- Fuel remaining
- Battery level remaining
- Air quality

Subjective

- Was successful mission completion impacted by a system issue? (expert evaluation)
- Was appropriate consideration given to submarine systems? (expert evaluation)
- Was safety ever threatened by a submarine system fault? (subject response to be evaluated by expert)
- Was covertness ever threatened by a submarine system fault? (subject response to be evaluated by expert)

B.2 Summary Slides

Summary of Discussion Areas

- Communications (external) x 2
- Shared SA (includes internal crew interaction) x 3
- Safety x 2
- Ship Handling (formerly Seamanship) x 2
- Covertness x 3
- Planning x 2
- Contact Management x 2
- Individual SA x 3
- Submarine Systems x 2

Key Focal Points

- Safety and covertness are guiding principles that influence how other 'constituent' work tasks are executed
  - Priority of goals 1) Safety of the sub and her personnel, 2) Remaining undetected, 3) Achieving the aim. Default hierarchy, but modified with CO intent and response to situations / environments
- Conditions under which measurements are taken as affected by vigilance and fatigue
  - Difficult to achieve through experimentation, but acknowledged as key influencer of performance
- Elegance of planning could be assessed by the number of concurrent activities able to do at once; gain as much information available in a given time window
  - Bilges, communications, snort, all sensor search, EW info, update weather picture rain, wind, projection of how it affects you
  - Tools that may facilitate the speed at which you can do things would allow the experimenter to do a comparative study
- Sit reps are a stream of consciousness that has juicy information, but are difficult to assess
  - Provides experimenters insight into appropriate level of comms required to achieve optimal SA among crew
- Subjective measures a tough nut to crack given nature of assessment
  - Real time probe questions provide some insight, and may not necessarily be intrusive given operators training to focus on primary task
  - Intra and inter-rater reliability difficult to maintain

Key Focal Points

- Scenario profile needs to be tailored for the MOP or MOE under consideration (event driven aspects that necessitate the attention to specific components
  - E.g., Covertness is not always critical when no threat results from being detected
- Validation of measurement areas - no obvious objections from SMEs, but need to delve down further into the key metrics which are both demonstrable (through testing) and achievable through simulator development
- General necessity for emphasis on Fire Control Solutions as well as TMA (validity)
  - Importance of mental math, helps expedite solutions when needed (e.g., return to PD quickly)
  - Possibility of complacency, reliance on computed solutions
  - Fundamental understanding of the key pieces of information is important; not just ‘stacking up the dots’
- Planning and ship handling wrt maintaining / re-gaining tactical advantage
  - Deviation from course can often be acceptable when traded off with tactical disposition
  - E.g., appropriate turning to check for blind areas, utilization of environment (land),
- Submarine systems – a good schematic representation
  - Makes the job significatnly easier in terms of SA, particularly translating any faults into an understanding of what the ship can or cannot do when

List of symbols/abbreviations/acronyms/initialisms

AIS	Automatic Information System
AOB	Angle On the Bow
C2	Command and Control
CAE	Canadian Aviation Electronics
CO	Commanding Officer
CF	Canadian Forces
COI	Contact Of Interest
CPA	Closest Point of Approach
CWA	Cognitive Work Analysis
DND	Department of National Defence
DRDC	Defence Research and Development Canada
EMT	Eye Movement Tracking
GDTA	Goal Directed Work Analysis
HF	Human Factors
ID	Identification
IID	Integrated Information Display
IMO	International Maritime Organization
MOE	Measure of Effectiveness
MOP	Measure of Performance
MPA	Maritime Patrol Aircraft
NASA TLX	National Aeronautics and Space Administration Task Load Index
NATO	North Atlantic Treaty Organization
OOW	Officer Of the Watch
OBJ	Objective
OPI	Office of Primary Interest
ORBAT	Order Of Battle
PD	Periscope Depth
R&D	Research & Development
SA	Situation Awareness
SIT REP	Situation Report

SME	Subject Matter Expert
SUBJ	Subjective
SUBOPAUTH	Submarine Operating Authority
TG	Task Group
TMA	Target Motion Analysis
vVic	Virtual Victoria

This page intentionally left blank.

Distribution list

Document No.: DRDC Atlantic CR 2011-282

LIST PART 1: Internal Distribution by Centre:

1	Francine Desharnais
1	Mark Hazen
1	Aren Hunter
1	Jacqui Crebolder
1	Tania Randall
3	DRDC Atlantic Library
8	TOTAL LIST PART 1

LIST PART 2: External Distribution by DRDKIM

1	Library and Archives of Canada, Attn: Military Archivist, Government Records Branch
1	DRDKIM
2	TOTAL LIST PART 2

10 TOTAL COPIES REQUIRED

This page intentionally left blank.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g., Centre sponsoring a contractor's report, or tasking agency, are entered in Section 8.)

   CAE Professional Services (Canada) Inc.,
   1135 Innovation Dr., Ottawa, ON K2K 3G7

2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.)

   Unclassified
   (NON-CONTROLLED GOODS)
   DMC A
   REVIEW: GCEC JUNE 2010

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C, R or U) in parentheses after the title.)

   Development of Measures of Effectiveness and Performance from Cognitive Work Analysis Products

4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used)

   Lai, G.; Lamoureux, T.

5. DATE OF PUBLICATION (Month and year of publication of document.)

   February 2012

6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 100

6b. NO. OF REFS (Total cited in document.)

   10

7. DESCRIPTIVE NOTES (The category of the document, e.g., technical report, technical note or memorandum. If appropriate, enter the type of report, e.g., interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

   Contract Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

   Defence R&D Canada – Atlantic

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) 11BA

9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

   W7707-098218/001/HAL

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)

   5254-001

10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

   DRDC Atlantic CR 2011-282

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

   Unlimited

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

   Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

Defence Research and Development Canada (DRDC) – Atlantic is currently investigating new systems to support the VICTORIA Class Submarine command team. The program will include human-in-the-loop experimentation within a virtual environment facility known as vVictoria. The project had previously utilized human factors engineering processes such as Cognitive Work Analysis (CWA) to characterize the command team's work environment in support of new system design. This report investigated the utility of re-purposing the results of the design-based work to assist in developing measures of effectiveness (MOE) and performance (MOP) to support project experimentation. It was found that while the design work provided a lot of information, specific augmentation for measure development was required. Further, it was found that the addition of specific mission vignettes was required to reduce the numbers of measures coming out of the design work to a manageable level. This report details the development process, results of extra knowledge elicitation and validation, and resulting MOE/MOPs.

Recherche et développement pour la défense Canada (RDDC) – Atlantique examine actuellement de nouveaux systèmes servant à soutenir l'équipe de commandement des sous-marins de la classe VICTORIA. Ce programme comprendra des essais avec intervention humaine dans le simulateur virtuel vVictoria. Dans le cadre de ce projet, des processus d'ingénierie des facteurs humains tels que l'analyse du travail cognitif (ATC) ont servi à caractériser le milieu de travail de l'équipe de commandement afin d'appuyer le nouveau concept du système. Ce rapport vise à déterminer si les résultats du travail de conception peuvent aider à élaborer des mesures de l'efficacité (MOE) et du rendement (MOP) en vue de réaliser les essais du projet. Beaucoup d'information a été recueillie durant le travail de conception, mais il faut davantage de renseignements spécifiques pour établir les mesures. En outre, d'autres scénarios de mission spécifique sont nécessaires pour amener le nombre de mesures obtenues avec le travail de conception à un niveau raisonnable. Le présent rapport porte sur le processus d'élaboration, les résultats liés à l'acquisition et la validation d'autres connaissances, ainsi que les MOE et les MOP qui en découlent.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g., Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

CWA; MOE; Submarine

Defence R&D Canada

Canada's leader in defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE R&D DÉFENSE

www.drdc-rddc.gc.ca